

## 4 Aspekte fokussierter räumlicher Aufmerksamkeit

„[...] a useful, indeed perhaps a *necessary* strategy for understanding human behavior is first to try to discover some properties of the mechanisms which operate on knowledge and which bring the organism in contact with its environment“ (Pylyshyn 1994:364)

### 4.1 Vorwort: Warum eine Betrachtung visueller Aufmerksamkeit?

„While orienting to stimuli in visual space is a restricted sense of attention, I believe that its study is capable of providing us both with important tests of the adequacy of general models of human cognition and with new insights into the role of attention in more complex human activity.“ (Posner, 1980:4)

Wie läßt sich die Frage nach den Eigenschaften räumlichen Wissens sowie nach der Beziehung von Sprache und Raum beantworten, ohne in den interdisziplinären Zirkel der Charakterisierung räumlicher Relationen zu geraten und ohne die Unterscheidung impliziter und expliziter räumlicher Relationen zu mißachten? Ein Zugang zur Beantwortung dieser Frage, der in dieser Arbeit verfolgt wird, besteht darin, die Bildung räumlichen Wissens einerseits unter dem Gesichtspunkt der notwendigen *Reduktion* durch *Selektion* von Information zu betrachten, die beim Aufbau mentaler Repräsentationen stattfindet, und andererseits unter dem Gesichtspunkt der *Konstruktion* dieser Repräsentationen. Weder können *alle* Informationen repräsentiert, noch können sie alle *gleichzeitig* verarbeitet und für die Auswahl verhaltensrelevanter Aktionen verwendet werden.<sup>1</sup> Zudem existiert keine vorgefertigte Welt, stattdessen müssen Interpretationen erst hergestellt (vgl. zum Beispiel die Interpretation von Vexierbildern) und repräsentationelle Elemente gebildet werden (s. Beobachtung B3). Der Aspekt der Reduktion selbst bedarf keiner weiteren Begründung und wird hier nicht weiter thematisiert. Die Aspekte der Selektion und Konstruktion sind hingegen zentral für die vorliegende Arbeit. Sie sind eng mit dem Begriff der Aufmerksamkeit verknüpft (hier überwiegend eingeschränkt auf visuell-räumliche Aufmerksamkeit), von dem gezeigt werden soll, daß er sowohl einen inhärenten Zusammenhang zu kognitiv-räumlichen als auch zu sprachlich räumlichen Relationen aufweist.

Selbstverständlich nehmen wir unsere Umwelt visuell allein deswegen ausschnittsweise wahr, weil in einem beschränkten Zeitraum nur eine beschränkte Menge an Informationen

<sup>1</sup> „The visual system cannot fully process all of its input. There is not enough room in the skull for all of the neural hardware that would be required to perform all visual functions at all locations in the visual field at the same time [...] The visual system has two basic approaches to this problem. The first is to discard input. [...] The second approach is to process information selectively“ (Wolfe 1994:202).

durch die Augen geliefert wird (visuelles Feld). Experimente im Bereich visueller Aufmerksamkeit haben jedoch gezeigt, daß gerade auch bzgl. dieser Menge an Informationen Selektion stattfindet. Dies läßt sich anhand von Experimenten nachweisen, die im Paradigma der visuellen Suche durchgeführt wurden (Theeuwes 1993). Präsentiert man Versuchspersonen ein visuelles Feld, in dem ein einzelnes Element (Target) in einer Menge gleichartiger anderer Elemente (Distraktoren) gefunden werden soll, so ist diese Suche subjektiv einfach und – unabhängig von der Anzahl der Distraktoren – (relativ) gleichbleibend schnell (sogenanntes „Popout“-Phänomen). Ist das Target hingegen schwer von den Distraktoren unterscheidbar<sup>2</sup> (oder gar nicht vorhanden), so steigt die Reaktionszeit der Verifikation mit der Anzahl der Distraktoren. Diese Ergebnisse haben zu einer Unterscheidung zweier Verarbeitungstypen bzw. zweier Verarbeitungsebenen geführt: eine Ebene *paralleler* Verarbeitung von Merkmalen des visuellen Feldes einerseits und eine Ebene *serieller* Verarbeitung (bei schwierigen Diskriminationen oder komplexen Elementen) andererseits. *Visueller Aufmerksamkeit* wird dabei im allgemeinen die Rolle der selektiven Komponente zugesprochen, die die beschränkte und lokal begrenzte serielle Verarbeitung kontrolliert.

Selektive Aufmerksamkeit ist offenbar außerdem mit der Konstruktion repräsentationeller Elemente assoziiert. Nach Kahneman/Treisman (1992) führt die Fokussierung von Aufmerksamkeit auf ein neues Element zur Bildung visuell-räumlich adressierter *object files*, die die subjektive Objektkonstanz perzeptueller Objekte gewährleisten, obwohl diese sowohl im Hinblick auf ihre Merkmale als auch auf ihre Typzuordnung erheblichen Änderungen ausgesetzt sein können:

„Onlookers in the movie can exclaim 'It's a bird; it's a plane; it's superman!' without any change of referent for the pronoun. If the appropriate constraints of spatiotemporal continuity are observed, objects retain their perceptual integrity and unity. Since neither spatial location, sensory properties, nor even the most appropriate label need remain constant, we are forced to attribute any object-specific perceptual phenomena to some form of object-specific representation, addressed by its present location and by its continuous history of travel and change through space over time“ (Kahneman/Treisman 1992:217).

Aufmerksamkeit wird außerdem als notwendig für die Integration einzelner Merkmale zu einem Objekt angesehen (Treisman 1988) und übt einen wesentlichen Einfluß auf den Interpretationsprozess ambiguer perzeptueller Konstellationen (Peterson/Gibson 1991), den Übergang von orientierungsabhängigen zu orientierungsunabhängigen Repräsentationen bei der Identifikation disorientierter Objekte (Murray 1995) sowie die Modulation globaler vs. lokaler Verarbeitung hierarchisch strukturierter Objekte (Kosslyn 1994:96f) aus.

Ein weiterer wesentlicher Aspekt, der den Kern der in dieser Arbeit geführten Argumentation darstellt, ist die Erkenntnis, daß visuelle räumliche Aufmerksamkeit nicht nur stimulus-abhängig („bottom-up“) geleitet ist, sondern auch top-down gesteuert operieren kann. Üblicherweise wird hierfür eine entsprechende Komponente des kognitiven Systems angenommen („Aufmerksamkeitskontrolle“). Will man jedoch nicht einem Homunkulus-Fehlschluß anheimfallen, so muß diese Komponente grundsätzlich repräsentationsbasiert (d.h. letztendlich als (auch) durch den Input charakterisiert) konzipiert werden. Solche Repräsentationen bestimmter Aspekte der *Wahrnehmung von Raum* können – obgleich nur

<sup>2</sup> Die Schwierigkeiten können dabei darauf zurückzuführen sein, daß Target und Distraktor gemeinsame Merkmale aufweisen oder daß das Target ein Merkmal entbehrt, das die Distraktoren aufweisen.

indirekt und somit nicht als Morphismen erster Ordnung – auch als Repräsentationen von Raum aufgefaßt werden. Von solchen Repräsentationen soll gezeigt werden, daß sie unmittelbar der Repräsentation sprachlicher Raumkonzepte zugrundeliegen.

#### 4.2 Ein Wort der Vorsicht zum Thema „Aufmerksamkeit“

„Everyone knows what attention is. It is the taking possession by the mind, in clear and vivid form, of one out of what seems several simultaneously possible objects or trains of thought“. (William James)<sup>3</sup>

Grundsätzlich stellt sich einer Arbeit, die einen Terminus als erklärendes Konstrukt *verwenden* will, die Aufgabe, diesen Terminus zunächst hinreichend genau zu charakterisieren bzw. zu definieren und dann seine Relevanz für die Theoriebildung zu motivieren. Eine Arbeit, die in dieser Weise auf ‚Aufmerksamkeit‘ rekurriert, muß allerdings die folgenden Beobachtungen Johnston/Darks zur Kenntnis nehmen, die als die Spitze des Eisbergs der Problematik seiner Verwendung angesehen werden können:

„In reviewing the literature on attention we were struck by several observations. One was a widespread reluctance to define attention. Another was the ease with which competing theories can accommodate the same empirical phenomena. A third observation was the consistent appeal to some intelligent force or agent in explanations of attentional phenomena. [...] As a consequence, the more we read, the more bewildered we became“ (Johnston/Dark 1986:43).

Die Untersuchung von Aufmerksamkeit stellt sich somit selbst in dem hier eingeschränkten Bereich als äußerst heterogenes Forschungsgebiet dar, das durch die Fragen nach dem „Ort“ von Aufmerksamkeit und nach der Identifikation aufmerksamkeitsbestimmter Prozesse (gegenüber solchen, die ohne Aufmerksamkeit ablaufen) bestimmt ist. Trotzdem die experimentalpsychologische Forschung hierzu gerade in jüngerer Zeit durch Ergebnisse aus dem Bereich der Neuropsychologie/-physiologie (z. B. PET-Scans oder ERP-Aufnahmen<sup>4</sup>) oder durch pathologische Studien (z. B. die Neglekt-Forschung) ständig bereichert wird, ist eine umfassende bzw. einheitliche Theorie der Aufmerksamkeit noch nicht in Sicht und sollte nach Allport auch nicht angestrebt werden:

„The preeminence accorded to these two questions over the past twenty-five years rested [...] on the belief in the possibility of a unitary (and simple) 'theory of attention'. Indeed, answers to these questions were frequently put forward as just that—as a putative theory of attention. This belief rested, in turn, on the assumption (implicit or explicit) that attentional functions were *all of one type* [...]. The penalty for such wishful thinking is to be condemned forever to appeal, in one's theory, to ill-defined (or even completely undefined) causal mechanisms and constraints—attention, attentional resources, central processing system, (anterior) attentional system, central executive, further processing, and the like—whose explanatory horsepower is nil“ (Allport 1993:206).

<sup>3</sup> Zitiert in Posner/Raichle (1994:154).

<sup>4</sup> Positron Emission Tomography und Event Related Brain Potentials, vgl. Posner/Raichle (1994).

Allport zeigt allerdings einen Weg auf, diesen Mißstand zu vermeiden: „It is by taking seriously the idea that attentional functions are of very many different kinds, serving a great range of different computational purposes“ (ibid.).<sup>5</sup> In diesem Sinne und vor diesem Hintergrund sollen in diesem Kapitel Aufmerksamkeitsphänomene soweit beleuchtet und bestehende Theorien und Begrifflichkeiten soweit betrachtet werden, wie sie dazu beitragen, Aussagen über die hier relevanten Aspekte von ‚Aufmerksamkeit‘ (d. h. zur Repräsentation räumlicher Relationen beitragende Funktionen) zu ermöglichen.

Gerade weil Termini wie ‚Fokus‘, ‚Aufmerksamkeit‘, ‚Fokussierte Aufmerksamkeit‘ im allgemeinen sehr unterbestimmt sind und daher äußerst unterschiedlich verwendet werden,<sup>6</sup> soll an dieser Stelle betont werden, daß sie *nicht* von vornherein in irgendeiner Beziehung zu dem in dieser Arbeit verwendeten Terminus der ‚selektiven räumlichen Aufmerksamkeit‘ stehen (allerdings ist es sicherlich eine interessante Aufgabe, etwaige Beziehungen im Nachhinein aufzuzeigen). Ich halte es daher mit Kratzer, die ihre Ausführungen zu ‚focus‘ wie folgt einschränkt: „The category focus is notoriously obscure, and we will not try to do justice to the many usages that survive in the literature“ (Kratzer, op. cit., 804).

### 4.3 Aufmerksamkeit versus Augenbewegungen

„The basic idea is that when *sequentially* fixating different objects, the change in fixation provides a direct encoding of the desired spatial relationship“ (Ballard 1987:192).

Ohne nähere Kenntnis vorliegender Evidenz zu Aufmerksamkeitsphänomenen mag der Versuch, eine Verbindung von Aspekten visueller Aufmerksamkeit und Aspekten räumlicher Relationen zu etablieren, zunächst ein wenig erstaunen. Wenn schon eine Beziehung zwischen Raumrepräsentation und visueller Wahrnehmung hergestellt werden soll, wieso bildet nicht die Untersuchung von Augenbewegungen den primären Fokus des Interesses? Untersuchungen zeigen, daß Augenbewegungen in systematischer Weise dazu dienen, sukzessive die informativen Aspekte des betrachteten Bereichs zu erfassen:

„Early work on this topic has indicated that fixations are not distributed randomly over a picture, but rather that a relatively large proportion of the fixations is allocated to a relatively small portion of the scene [...]. A number of more recent studies have demonstrated that the gaze is attracted to ‚informative areas‘ of a picture, where informative is defined in terms of subjective ‚informativeness ratings‘ by independent observers“, Loftus/Mackworth 1978:565).

<sup>5</sup> vgl. auch Neumann (1992).

<sup>6</sup> Zu einem Überblick über den linguistischen ‚Fokus‘-Begriff s. z. B. A. Kratzer (1991), „Current Issues in the Theory of Focus“, in A. v. Stechow, D. Wunderlich (Hrsg.), *Semantik: ein internationales Handbuch der zeitgenössischen Forschung* (Berlin et al.: de Gruyter), 804-834; zu Verwendungen in der sprachorientierten Künstliche-Intelligenz-Forschung, Computerlinguistik, Psycholinguistik und Psychologie s. z.B. Grosz/Sidner 1986, McKeown 1985, Brennan 1995, Rinck/Bower 1995.

Relativ längere Phasen, in denen die Augen unbeweglich sind (→ Fixationen) und in denen offenbar die am fixierten Ort befindliche Information verarbeitet wird, werden dabei von kurzen Sprüngen der Augen (→ Sakkaden) getrennt. Die so entstehenden Augenbewegungspfade des Absuchens eines visuellen Bereichs („Scanpaths“) werden anscheinend außerdem repräsentiert und dienen darauffolgend dazu, einen leitenden top-down-Einfluß auf aktuelle Augenbewegungen auszuüben (Groner 1988). Obwohl also Augenbewegungen offenbar Information über räumliche Beziehungen herstellen (s. das Zitat von Ballard), gibt es trotzdem mindestens drei Einwände gegen die *primäre* Relevanz dieses Phänomenbereichs für die Untersuchung der Beziehung von Wahrnehmung und Raumrepräsentation.

Erstens wird zwar der Fokus der Verarbeitung durch Augenbewegungen verschoben, doch können räumliche Relationen auch ohne sie hergestellt werden („Directing the processing focus [...] may be achieved in part by moving the eyes [...]. But this is clearly insufficient: many relations [...] can be established without eye movements. A capacity to shift the processing focus internally is therefore required“, Ullman 1984:561). Zweitens müssen Augenbewegungen aufgrund prinzipieller theoretischer Überlegungen von räumlicher Verarbeitung getrennt werden („Eye movements are an important component of behavior in visual space, but spatial behavior must be dissociable from eye movements, for otherwise we could never attend to and act on objects in our peripheral retina without making an eye movement to them“, Goldberg 1982:284). Die Dissoziierbarkeit von Augenbewegungen und visuell-räumlicher Orientierung ist zudem empirisch anhand von Einzelzellableitungen im Gehirn von Affen (ibid.) und im Rahmen von Aufmerksamkeitsexperimenten (Posner 1980) nachgewiesen worden. Drittens stellt visuelle Wahrnehmung nur eine spezifische Modalität der Interaktion eines Subjekts mit seiner (räumlichen) Umwelt dar (neben der haptischen und auditiven). Aufmerksamkeit kann im Vergleich dazu als ein allgemeineres, intermodales Konstrukt angesehen werden, dem möglicherweise eine entsprechende Rolle in einem umfassenden räumlichen Repräsentationssystem, wie es von Bryant (1992) vorgeschlagen wird, zukommt.

#### 4.4 Der funktionale Aspekt: Selektivität der Aufmerksamkeit

##### 4.4.1 Selektion und Aufmerksamkeit

Es gibt mindestens drei grundlegende Auffassungen von ‚Aufmerksamkeit‘, die unterschieden werden können (vgl. Umiltà 1988). Unter Aufmerksamkeit kann zum einen ein Ressourcen-verwaltender Prozess verstanden werden, bei dem einer bestimmten Aufgabe Ressourcen zur Verfügung gestellt und dafür anderen Aufgaben entzogen werden. Zum anderen kann damit ein die Aufnahmefähigkeit regulierender Prozess gemeint sein, durch den z. B. die Wachsamkeit bzgl. sensorischer Informationen erhöht werden kann. Schließlich kann Aufmerksamkeit als ein selektiver Prozess verstanden werden, durch den bestimmte Informationen bewußter Verarbeitung zur Verfügung gestellt wird, so daß andere dafür nur unbewußt oder gar nicht verarbeitet („ausgefiltert“) wird.

Selektion selbst ist keinesfalls ein einheitliches Phänomen. Sie kann auf die Auswahl bestimmter modalitätsspezifischer (z. B. auditiver vs. visueller) oder positionsspezifischer (z. B. an einen bestimmten visuell-räumlichen „Ort“ gebundene) Information, oder auf die Auswahl von Elementen mit bestimmten Merkmalen (Farbe, Form etc.) bzw. von bestimmter Kategorienzugehörigkeit (Lebewesen, Artefakt, Mensch, Tier etc.) referieren. All diesen Erscheinungen ist gemeinsam, daß Information, die durch Aufmerksamkeitszuwendung „selegiert“ worden ist, besser verarbeitet wird. Damit ist allerdings noch nichts darüber ausgesagt, welche Mechanismen der Selektion zugrunde liegen (wann, wie und wo sie stattfindet). Selektion kann einerseits als durch eine *Verstärkung* der attendierten Information oder als durch *Filterung* der nicht-attendierten Information hervorgerufen betrachtet werden.<sup>7</sup>

Ebenfalls bleibt noch unklar, ob Selektion allgemein als Ursache (d. h. als spezifischer kausaler Auswahlmechanismus) oder als Effekt (d. h. Auswahl ist das Ergebnis nicht-spezifischer Verarbeitung) aufgefaßt werden soll.<sup>8</sup> Dies entspricht zum Teil der plausiblen scheinenden Frage, ob Aufmerksamkeit als Verwaltungsmechanismus einer begrenzten Ressource oder als die verwaltete Ressource selber zu verstehen ist. Der dabei präsupponierte Ressourcenbegriff wird allerdings von einigen Autoren zugunsten einer als ‚selection for action‘-Sichtweise infrage gestellt (Allport 1993, Neumann 1990).

#### 4.4.2 Ausrichtung („orienting“) der Aufmerksamkeit

Eine allgemeine Sicht der Funktionsweise selektiver Aufmerksamkeit sowie eine Reihe grundlegender Termini lassen sich auf Posner (1980) zurückführen. Er unterscheidet prinzipiell die Zuwendung oder Ausrichtung von Aufmerksamkeit auf eine bestimmte Struktur („orienting“) von der Tatsache, daß sich eine Person der attendierten Struktur bewußt ist („detecting“):

„I will use the term *orienting* to mean the alignment of attention with a source of sensory input or an internal semantic structure stored in memory. [...] I distinguish orienting from another cognitive act that I call *detecting*. By detecting I will mean that a stimulus has reached a level of the nervous system at which it is now possible for the subject to report its presence[...]“ (Posner 1980:4).

Weiterhin unterscheidet Posner zwischen *externer* (durch Stimuli determinierter) und *interner* (durch mental repräsentierte „search plans“ determinierter) Kontrolle des orienting-Prozesses sowie zwischen *offenen* („overt“; mit Kopf- und Augenbewegungen gekoppelten) und *verdeckten* („covert“; eigenständig ablaufenden) Aufmerksamkeitswechseln. Insbesondere weist er in seinen als „speeded detection“- bzw. „cueing“-Paradigma bekannt gewordenen Untersuchungen die Existenz von covert orienting nach. Hierzu mußten Versuchspersonen auf links oder rechts von einem zentralen Fixationspunkt erscheinende

<sup>7</sup> Die zweite Auffassung geht auf die „Filter-Theorie“ Broadbents zurück. In den meisten Ansätzen überwiegt jedoch die erste Auffassung, in einigen neueren wird – insbesondere zur Erklärung des ‚inhibition effect‘ (s. Umiltà 1988:189ff) bzw. ‚negative Priming‘ (s. Fox 1995) – auch eine gemischte, duale Sichtweise vorgeschlagen.

<sup>8</sup> Nach Johnston/Dark (1986) und Allport (1993) sind die meisten gegenwärtigen Theorien selektiver Aufmerksamkeit kausale Theorien. Entsprechend finden sich bei den Autoren Anspielungen auf das Homunculus-Problem.

Stimuli so schnell wie möglich reagieren. Vor jedem Target erschien als Cue an der zentralen Position entweder ein neutrales Symbol (ein ‚+‘, das die gleiche Wahrscheinlichkeit des Targets, links oder rechts aufzutauchen, anzeigte) oder ein Pfeil (bei dem das Target mit einer Wahrscheinlichkeit von 0.8 in der angegebenen Richtung erschien). In den Ergebnissen zeigte sich deutlich ein Vorteil von validen (Target erscheint an der angezeigten Position) gegenüber neutralen und invaliden Durchläufen. Wesentlich hierbei war, daß die Augenbewegungen kontrolliert und nur solche Durchläufe verwendet wurden, in denen die Augen fixiert blieben. Auf diese Weise wurde der Nachweis der Existenz von *covert orienting* erbracht.

Weitere Experimente in diesem Paradigma zeigten, daß Wissen über räumliche Position eines Targets bei komplexeren Aufgaben erst bei einem überfüllten visuellen Feld zum Tragen kommt, bei einfachen Aufgaben hingegen eher hinderlich ist: „when given a single key to press whenever they saw a digit, subjects could hardly avoid false alarms when a letter appeared at the correct spatial position“ (Posner 1980:8). Es zeigte sich weiterhin, daß orienting *nicht* als statisches Konstrukt (z. B. räumlich positionierbarer Filter) aufgefaßt werden darf: Wurde eine bestimmte Position für eine ganze Reihe von Durchläufen als sehr wahrscheinlich ausgezeichnet, so ergaben sich keine Vorteile dieser Position im Vergleich zu Bedingungen, in denen alle Positionen *gleich* wahrscheinlich waren.

Ein wesentliches Ergebnis von Posners Arbeit ist außerdem die Klärung der Beziehung von Aufmerksamkeit und Augenbewegungen: Einerseits läßt sich feststellen, daß bei durch einen peripheren Stimulus ausgelösten Augenbewegungen die Tendenz dazu besteht, daß Aufmerksamkeitszuwendung auf diesen Reiz *vor* den Bewegungen selbst stattfindet; andererseits lassen sich Bedingungen konstruieren, in denen sich Aufmerksamkeit und Augen entgegengesetzt bewegen. Nach Posner liegt somit (nur) eine funktionale Relation vor.

Posner findet schließlich Hinweise auf einen Unterschied zwischen der Verarbeitung peripher und zentral vorgegebener Cues („Pull“- vs. „Push“-Cues), die auf eine notwendige Differenzierung *exogener* (stimulusabhängiger) und *endogener* (zentraler) bzw. *automatischer* und *willentlicher* Kontrolle hindeuten (s. hierzu auch Briand/Klein 1987).

#### 4.4.3 „Frühe“ vs. „Späte“ Selektion

Es ist keineswegs von vornherein ersichtlich und entsprechend gegenwärtig noch nicht geklärt, *wann* Selektion im Verlauf der Verarbeitung visuellen Inputs stattfindet. Grundlegender Konsens besteht zunächst nur darüber, daß zwischen zwei relevanten Verarbeitungsebenen unterschieden werden muß (Theeuwes 1993): einer Ebene prä-attentiver Verarbeitung, die keine Kapazitätsbeschränkungen aufweist und auf der dementsprechend Merkmale des visuellen Feldes *parallel* verarbeitet werden können, und einer Ebene attentiver Verarbeitung, deren Ressourcen beschränkt sind und auf der nur ein Element (bzw. in jedem Fall nur wenige) gleichzeitig berücksichtigt werden kann, was eine *serielle* Verarbeitung notwendig macht. Diese Ebenen stehen in einer hierarchischen Beziehung zueinander, so daß „Selektion“ als Übergang eines Elements von der ersten zur zweiten aufgefaßt werden kann. Unstrittig ist ebenfalls, daß das *Entdecken* eines Elements im visuellen Feld und seine *Identifikation* als Element von einem bestimmten Typ systematisch getrennt werden müssen.

Uneinigkeit besteht allerdings darüber, welche Operationen attentiv durchgeführt werden. „Frühe“ Selektion entspricht der Auffassung, daß Merkmale des visuellen Feldes zwar bereits entdeckt sein können, aber bei attentiver Zuwendung noch nicht identifiziert sind (Sagi/Julesz 1985). Sie kann außerdem bedeuten, daß auf der attentiven Ebene noch perzeptuelle Operationen durchgeführt werden, z.B. die Integration visueller Merkmale zu einem Objekt (Treisman 1988, Kahneman/Treisman 1992).<sup>9</sup> Mit „später“ Selektion ist hingegen die Annahme verbunden, daß Aufmerksamkeitszuwendung keinen Einfluß mehr auf die perzeptuelle Verarbeitung hat und daß die Elemente des visuellen Feldes bereits parallel identifiziert worden sind. Aus diesem Grund sind Späte-Selektion-Ansätze eher *objekt-basiert* (Duncan 1984), während in Frühe-Selektion-Ansätzen der Bezug auf den visuellen *Raum* eine wesentliche Rolle spielt.

Die Möglichkeit, mehrere Objekte des visuellen Feldes gleichzeitig verfolgen zu können („visual tracking“, vgl. Yantis 1992), scheint dieser Zuordnung allerdings nicht zu entsprechen. In von Pylyshyn durchgeführten Experimenten (beschrieben in Pylyshyn 1994) mußten Versuchspersonen jeweils drei und sechs Targets in einer größeren Menge sich beliebig bewegender identischer anderer Objekte („+“-Zeichen) verfolgen. Hierzu wurden die Targets zunächst durch Aufleuchten identifiziert, bewegten sich daraufhin für eine Dauer von 7 bis 15 Sekunden. Nach diesem Intervall leuchtete ein Quadrat auf und die VPs mußten entscheiden, ob dieses an einem Target-Objekt, einem Distraktor-Objekt oder an einer anderen Stelle im Display geschehen war. Hierbei war die Erkennungsrate (obwohl mit steigender Anzahl der Targets abnehmend) außergewöhnlich hoch. Dieses Phänomen korrespondiert zu der als „Subitizing“ bekannten Erscheinung, daß die Anzahl einer kleinen Menge visueller Objekte sehr schnell bestimmt werden kann (Trick/Pylyshyn 1993). Pylyshyn deutet diese Ergebnisse dahingehend, daß bereits auf der präattentiven Ebene eine bestimmte Anzahl von Objekten räumlich indiziert werden kann (durch sogenannte „Fingers of Instantiation“, FINSTs).

Bundesen (1990) nimmt eine Zwischenposition bzgl. der Früh/Spät-Dichotomie ein, die er gemäß seiner „Selektion ist perzeptuelle Kategorisierung eines Elements  $x$  zu einer Kategorie  $I$ “-Sicht folgendermaßen charakterisiert:

„In agreement with late-selection theories, it is assumed that strengths of sensory evidence for perceptual categorizations of the form " $x$  belongs to  $i$ " are computed before selection takes place; because  $i$  need not be defined by a simple physical feature, intelligent selection is possible. On the other hand, in agreement with early-selection theories, the recognition problem is resolved for no elements except those that are selected“ (Bundesen 1990:527).

<sup>9</sup> Nach der FIT-Theorie Treismans dient Aufmerksamkeit insbesondere dazu, durch eine *Konjunktion* von Merkmalen definierte Objekte erkennen zu können. Sagi/Julesz zeigen, daß auch für die Identifizierung einfacher Merkmale Aufmerksamkeitszuwendung notwendig ist.

## 4.5 Der qualitative Aspekt: Raum-basierte vs. Objekt-basierte Ansätze fokussierter Aufmerksamkeit

### 4.5.1 Metaphern raum-basierter Aufmerksamkeit

Motiviert durch die Auffassung einer lokal begrenzten und verstärkend wirkenden Zuwendung von Aufmerksamkeit auf Information, die sich an einer bestimmten visuell-räumlichen Position befindet, ist die Vorstellung, daß visuell-räumliche Aufmerksamkeit wie ein Lichtspot auf das visuelle Feld scheint. Diese Metapher verkörpert mehrere angenommene Eigenschaften visuell-räumlicher Aufmerksamkeit.

Zum einen ist sie primär und inhärent räumlich: Objekte oder Aspekte von Objekten werden stets *an einem bestimmten Ort* „erfaßt“. Diese Vorrangigkeit der Lokationseigenschaft von Objekten wird von Johnston/ Dark konstatiert und gleichzeitig anhand von Untersuchungen belegt: „[...] location proved to be more fundamental than color and shape. That is, subjects could not know the color or the shape of an item without knowing its location, but they could know its color without knowing its shape or vice versa.“ (Johnston/Dark 1986:49f, vgl. auch Tsal/Lavie 1993)). Objekte werden entsprechend der Metapher erst dann „erkannt“, wenn das Licht auf sie gerichtet ist.

Außerdem ist der Lichtspot (scharf?) begrenzt, zusammenhängend und von einer bestimmten Form und Größe. Dies spiegelt die Beobachtung wider, daß Aufmerksamkeit nicht gleichzeitig auf mehrere „Orte“ gerichtet sein kann (s. aber Castiello/Umiltà 1992).

Schließlich muß der Lichtspot erst auf eine bestimmte Stelle *gerichtet* werden, bevor die dort befindlichen Objekte „zur Kenntnis genommen“ werden können, was direkt dem Konzept des ‚orienting‘ entspricht. Allerdings stellt sich die Frage, ob, in der Terminologie der Metapher ausgedrückt, das Licht an einem Ort aus- und an einem anderen wieder angeschaltet wird (diskrete Aufmerksamkeitsverschiebung), oder ob sich der Lichtstrahl kontinuierlich über das visuelle Feld bewegt (analogische Aufmerksamkeitsverschiebung). Einige Untersuchungen scheinen letzteres zu bestätigen (vgl. Umiltà 1988:184f), wobei allerdings sowohl Hinweise auf Aufmerksamkeitsbewegungen in konstanter Zeit (bei variierender Geschwindigkeit) als auch auf Bewegungen mit konstanter Geschwindigkeit (in variierender Zeit) gefunden wurden.

Im Kontrast dazu weisen Ergebnisse aus dem *rapid serial visual presentation attention shift paradigm* (Reeves/Sperling 1986) bzw. der allgemeineren episodischen Theorie räumlicher Aufmerksamkeit (Sperling/Weichselgartner 1995) auf einen diskreten Wechsel der Aufmerksamkeitszuwendung hin (s. auch Chastain 1991) und sind entsprechend als das Öffnen bzw. Schließen von Toren („attention gating“) für den Zugang zu Information an einem bestimmten Ort konzeptualisiert worden.

Die Vorstellung des attentionalen Lichtspot als von einer bestimmten, fixen Größe ist in den Arbeiten von Eriksen (z. B. Eriksen/Yeh 1985, Eriksen/James 1986) zugunsten der sogenannten „Zoom-lens“-Metapher aufgegeben worden. Danach operiert Aufmerksamkeit nicht in einem einzigen ‚spotlight‘-Modus, sondern kann sowohl *verteilt* als auch *konzentriert (fokussiert)* sein, mit der folgenden Qualifikation:

„Although Jonides (1983) has distinguished these as two processes or modes of operation of visual attention, we do not believe they should be considered as two alternative modes, but rather as

two poles on a continuous range of attentional capacity distribution in the visual field“ (Eriksen/Yeh 1985:595).

Die Größe des Aufmerksamkeitsbereichs hängt von den Aufgabenanforderungen ab. Je nachdem, wieviel Information z. B. zum Erreichen einer bestimmten Entscheidung benötigt wird, vergrößert oder verkleinert sich der Aufmerksamkeitsbereich. Dabei wird von einer gleichbleibenden Menge verteilter Ressourcen ausgegangen, so daß die Ausdehnung des Bereichs und die Menge der für die an einem bestimmten Punkt zur Verfügung stehende Energie (die „Auflösung“ der Linse) in einer reziproken Relation zueinander stehen. Entsprechend können bei distribuiertem Aufmerksamkeitszuwendung zwar mehr Objekte verarbeitet werden, dafür aber mit schlechterer Qualität. Die Fokussierung von Aufmerksamkeit schränkt hingegen zwar die Anzahl der betrachteten Objekte ein, verbessert jedoch die Informationsverarbeitung.

In dem *Gradientenmodell* LaBerge/Browns (1989) wird die Spotlight-Metapher zugunsten eines komplexeren Modells der Interaktion unterschiedlicher Domänen aufgegeben. Aufmerksamkeitszuwendung stellt sich in diesem Modell als das Öffnen eines Kanals („channels“) zwischen einer Domäne der Merkmalsregistrierung und einer Filterdomäne dar. Die Verteilung von Ressourcen ist dabei nicht durch die „Öffnungsweite“ des Kanals, sondern durch einen räumlich organisierten Aktivationsgradienten einer dritten Domäne („location expectation domain“) bestimmt (s. Abb. 36), der somit einem inhomogen leuchtenden und unscharf begrenzten Lichtfleck entspricht.

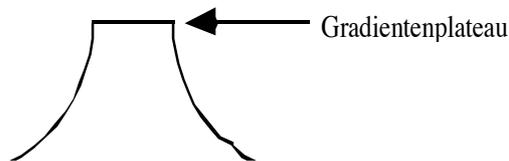


Abb. 36: Gradient der Aufmerksamkeit nach LaBerge/Brown

Aus Sicht der Autoren macht diese Unterscheidung die Annahme einer (kontinuierlichen) Bewegung des Aufmerksamkeitsbereichs in zweierlei Hinsicht überflüssig: Einerseits können Reaktionszeitunterschiede, wie sie LaBerge/Brown für an verschiedenen Orten innerhalb des Gradienten, aber nicht im Bereich des Gradientenplateaus erscheinende Targets fanden, durch die zum Rand hin abnehmenden Ressourcen erklärt werden. Mit steigender Entfernung vom Zentrum nimmt die Energie, die für das Öffnen eines Kanals bzw. für das Filtern eines Targets benötigt wird, ab („[...] nothing is moved across the visual field, but [...] a distribution of processing resources across the field affects the rate at which filtering will take place at a particular location.“ (LaBerge/Brown 1989:108). Andererseits können experimentelle Ergebnisse wie die von Posner (→ der Einfluß zentraler Cues auf das Erkennen peripherer Targets) durch die unterschiedliche Begünstigung der Etablierung eines sekundären Gradienten erklärt werden („the present theory accounts for the advantage of location precuing on reaction time by assuming that what the precue does is to provide time for a gradient to develop, not time for a channel to move“, *ibid*:119).

Aus den Unterschieden der skizzierten Modelltypen lassen sich die jeweils assoziierten Funktionen räumlich-basierter Aufmerksamkeit ableiten. Während auf der einen Seite (bei „gating“ und „channeling“-Modellen) der Aspekt des Filterns, d.h. das Unterdrücken unerwünschter und „Durchlassen“ erwünschter Information, im Vordergrund steht, wird auf der anderen Seite der Aspekt der Selektion durch Verstärkung an einem bestimmten Ort „sichtbarer“ Information betont.<sup>10</sup> Im zweiten Fall kann damit sowohl eine neuronale Verstärkung („Spatial orienting of visual attention is accompanied by *enhanced* neuronal responsiveness to visual stimuli appearing in the attended location“, Allport 1993:193) als auch eine allgemein verbesserte Verarbeitung attendierter Information („Attention can be likened to a spotlight that enhances the efficiency of detection of events within its beam“, Posner zitiert in Briand/Klein 1987:229) verstanden werden.

In der „Feature Integration Theory“ (FIT) Treismans (z.B. Treisman 1988) wird eine andere Funktion räumlicher Aufmerksamkeit betont. Aufmerksamkeit dient hier dem „Verkleben“ („Glueing“) separat verarbeiteter visueller Merkmale zu einem Objekt. (s. u., vgl. hierzu auch Briand/Klein 1987).

#### 4.5.2 Modelle Objekt-basierter Aufmerksamkeit

Objekt-basierte Aufmerksamkeitsmodelle gehen davon aus, daß Aufmerksamkeit nicht über einer „map of locations“ (Treisman) operiert, sondern präattentiv identifizierten Objekten zugewandt wird. Diese Hypothese wird durch zwei verschiedene Experiment-Typen gestützt. Der eine basiert im Kern auf der Annahme, daß es grundsätzlich schwierig ist, Aufmerksamkeit gleichzeitig auf verschiedene Objekte zu richten und daß das Reaktionszeitverhalten von VPs bei Aufgaben bzgl. zweier beliebiger Aspekte des visuellen Feldes davon abhängen sollte, ob jeweils nur ein oder aber zwei Objekte betroffen sind. Der relevante Kontrast ergibt sich dadurch, daß im ersten Fall (bei fokussierter Aufmerksamkeitszuwendung auf ein Objekt) beide Aspekte gleichzeitig verfügbar sind, während im zweiten Fall ein zeitaufwendiger Wechsel der Aufmerksamkeit notwendig ist.

Der andere Experiment-Typ ist durch das Bemühen gekennzeichnet, den Einfluß perzeptueller Gruppierung auf Target-Distraktor-Beziehungen nachzuweisen: Je nachdem, ob Target und Distraktor als zu unterschiedlichen perzeptuellen Objekten zugehörig oder nicht angesehen werden, sollten sich Unterschiede im Antwortverhalten zeigen.

In bezug auf den ersten Experiment-Typ führte Duncan (1984) eine Reihe von Experimenten durch, in denen VPs Displays wie in Abb. 37 gezeigt wurden, wobei sie auf jeweils zwei Aspekte (Größe des Rechtecks, Ort der Lücke, Neigung der Linie, Textur der Linie) reagieren mußten. Tatsächlich ergab sich ein signifikanter Unterschied in der Korrektheit der Reaktionen, ob z.B. nach Neigung und Schraffur der Linie, oder aber nach Größe des Rechtecks und Schraffur der Linie gefragt wurde.

<sup>10</sup> Die ebenfalls verwendete „Window“-Analogie (s. z.B. Kosslyn 1994) stellt eine Synthese dieser beiden Aspekte dar („Durchsichtigkeit“) und wird aus diesem Grund von Treisman/Sato (1990) favorisiert.

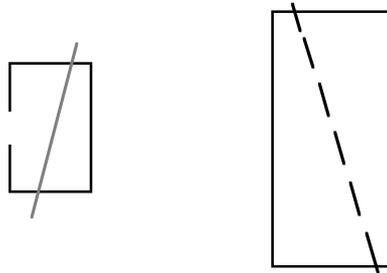


Abb. 37: Displays zum Test objektbasierter Aufmerksamkeit (Duncan 1984)

Duncan schlägt zwei mögliche Erklärungsansätze vor, um den beobachteten Phänomenen gerecht zu werden. Im ersten werden die unterschiedlichen Antwortzeiten auf die Beziehungen der beteiligten Aspekte entlang eines Kontinuums der Gruppierungsstärke (grouping strength) zurückgeführt. Innerobjektbeziehungen weisen danach eine größere Gruppierungsstärke auf als Zwischenobjektbeziehungen (vgl. Kramer/Jacobson 1991). Der zweite Erklärungsansatz basiert auf der Annahme einer hierarchischen Organisation visueller Informationen. Danach sind die Eigenschaften einer *Gruppe* von Objekten von den Eigenschaften einzelner Objekte<sup>11</sup> prinzipiell zu trennen, wobei der Wechsel fokussierter Aufmerksamkeit innerhalb einer Gruppe bestimmte Effekte (Interferenzen) zur Folge hat. Umgekehrt sind bei fokussierter Aufmerksamkeit alle Eigenschaften eines Objekts (auf seiner entsprechenden Ebene) verfügbar.

Die Annahme einer hierarchischen Informationsorganisation, insbesondere der hierarchischen Kodierung von Positionsinformation, wird von Baylis/Driver (1993) aufgegriffen. Sie unterscheiden *szenen-basierte* Kodierung von Rauminformation, in der die Position von Objekten zueinander festgehalten wird, von *objekt-basierter* Kodierung von Rauminformation, in denen die relative Position von Objektteilen repräsentiert ist.<sup>12</sup> Der Erklärungsansatz der Gruppierungsstärke wird hingegen in Kramer/Jacobson (1991) favorisiert. Sie schlagen als Alternative zu rein objekt-basierten Modellen ein Modell der Gruppierungsstärke vor, in dem räumliche Distanz einen wesentlichen Einfluß auf die Organisation perzeptueller Elemente ausübt:

„In the object-based model, the output is treated as a set of discrete groups of objects, whereas in the grouping-strength model, the output is treated in terms of elements organized along a continuum of grouping strength“ (Kramer/Jacobson 1991:282).

<sup>11</sup> Wobei auch einzelne Objekte rein räumlich als Gruppen von Teilobjekten aufgefaßt werden können: „Although the claim is that focal attention deals with whole objects, common sense strongly suggests that attention to a whole skyscraper would be inappropriate for determining whether there is a crack in the third window from the left on the 13th floor. What we need is an idea of the sorts of information likely to be coded in the ‚chunks‘ corresponding to ‚objects‘ at different levels of the hierarchy“ (Duncan 1984:514).

<sup>12</sup> Die Autoren ordnen in diesem Zusammenhang den Begriffen „szenen-basiert“ und „objekt-basiert“ die Begriffe „where“ und „what“ zu. Sie betonen dabei, daß die Identität eines Objekts wesentlich durch die Position seiner Teile zueinander bestimmt ist. What- und where-Informationen sind in dieser Sichtweise also nicht Typ-verschieden, sondern stellen nur unterschiedliche Ebenen der Repräsentation bzw. Verarbeitung visueller Information dar.

#### 4.5.3 Synthese Raum- und Objekt-basierter Modelle

Vecera/Farah (1994) weisen darauf hin, daß logisch nicht zwischen zwei, sondern zwischen drei Mechanismen visueller Aufmerksamkeit unterschieden werden muß. Aufmerksamkeit kann möglicherweise rein räumlich operieren (wobei Objekt-Aspekte *keine* Rolle spielen) oder ausschließlich objekt-basiert (wobei *nur* positionsinvariante Objekt-Aspekte, z.B. Formrepräsentationen, relevant sind). Im Einklang mit grouping strength-Modellen ist es aber auch vorstellbar, daß visuelle Objekte als gruppierte Lokationen betrachtet werden können, so daß die beobachtbaren Aufmerksamkeitseffekte gleichzeitig als objekt- und raum-basiert charakterisierbar sind.

Mit ihren Experimenten zeigten Vecera/Farah zunächst, daß sowohl rein objekt-basierte als auch auf gruppierten Lokationen basierende Mechanismen nachgewiesen werden können. Hierzu replizierten sie einerseits Duncans Versuchsreihe und fanden Unterschiede zwischen Ein- und Zwei-Objekt-Bedingungen, die unabhängig von räumlichen Aspekten (der Distanz zwischen den Objekten) waren. Dabei konnte eine Erklärung durch gruppierte Lokationen ausgeschlossen werden. Andererseits verwendeten sie in weiteren Experimenten *dieselben* Stimuli für Posner'sche Cuing-Aufgaben, in denen die Versuchspersonen auf den Onset eines Targets reagieren mußten, nachdem eines der Objekte als (valider oder invalider) Cue hervorgehoben worden war. Hier zeigte sich ein deutlicher Einfluß der Distanz zwischen den Objekten. Vecera/Farah folgern daraus, daß sowohl objekt- als auch lokations-basierte Aufmerksamkeitsmechanismen existieren, deren Anwendung jeweils von den aktuellen Anforderungen (einer Aufgabe) abhängt.

Die Diskussion um verschiedene Mechanismen der Aufmerksamkeit hat eine Parallele in einem eng verwandten Forschungsgebiet, das sich mit der „Inhibition of Return“ (IOR) von Aufmerksamkeit beschäftigt. Informell drückt IOR zunächst die Überlegung aus, daß visuelle Aufmerksamkeit gerade im Hinblick auf die Effizienz serieller visueller Suche nach Objekten möglichst redundanzfrei operieren sollte. D. h., um effektiv zu sein, sollte Aufmerksamkeit stets kontinuierlich auf neue Orte gerichtet werden und nicht häufig bereits attendierte Lokationen erneut „aufsuchen“. Anhand von Reaktionszeiten innerhalb des Cuing-Paradigmas konnten Posner und Kollegen dieses Phänomen (d.h. einen Mechanismus mit entsprechender Verhaltensrelevanz) tatsächlich nachweisen: Während Reaktionen auf eine gereizte *Position* („cued location“) anfänglich schneller sind als auf nicht gereizte Positionen (erklärt durch die Zuwendung von Aufmerksamkeit auf diesen Ort), so kehrt sich dieses Verhältnis ab einem Intervall von 300ms um und es stellt sich ein inhibitorischer Effekt ein.

In Analogie zu raum-basierten Aufmerksamkeitsmodellen schien IOR zunächst ausschließlich über Umgebungskoordinaten definiert („environment-based“) zu sein. Wird bei drei horizontal ausgerichteten Kästchen zunächst ein peripheres und danach wieder das zentrale gereizt, so zeigt sich, daß ab 300ms nach dem Onset des peripheren Reizes Antworten auf Targets in dem gereizten Kästchen langsamer sind als solche im ungereizten. Tipper und Kollegen (s. z.B. Tipper et al. 1994) wandelten dieses Experimentaldesign allerdings dahingehend ab, daß die peripheren Kästchen nach dem cuing um 90° gedreht wurden. Obwohl beide peripheren Kästchen jetzt gleichweit von der gereizten *Stelle* entfernt waren, ergab sich ein Inhibitionseffekt für das gereizte *Kästchen*, der in weiteren Versuchen erhärtet wurde und somit auf ein objekt-basiertes IOR hinweist. Tipper et al. Inter-

pretieren ihre Ergebnisse so, daß beide Mechanismen gleichzeitig existieren (abgeleitet aus Versuchen mit 180° Rotation der Kästchen) und in einer Hierarchie organisiert sind (mit dem environment-basierten Mechanismus als dem basalen und dem objekt-basierten als dem zusätzlichen, der insbesondere bei dem Verfolgen bewegter Objekte benötigt wird).

In Entsprechung zu der von Vecera/Farah angesprochenen Dreiteilung von Aufmerksamkeitsmodellen stellen Gibson/Egeth (1994) die Dichotomie von environment- und objekt-basiertem IOR infrage. Sie argumentieren – ähnlich wie Baylis/Driver (1993) –, daß „objekt-basiert“ nicht a priori mit „raum-invariant/-unabhängig“ gleichzusetzen ist. Stattdessen weisen sie darauf hin, daß Objekte auch als eine „Mikroumgebung“ aufgefaßt werden können, die durch die Repräsentation der objektkonstitutiven Lokationen gebildet wird.

#### 4.5.4 Perzeptuelle Gruppen und Hierarchien

Einer der Bereiche, in denen das Verhältnis von qualitativ unterschiedlichen Mechanismen der visuell-räumlichen Aufmerksamkeit einerseits sowie Hierarchien in der perzeptuellen Organisation andererseits explizit untersucht wird, ist das durch das Global/Lokal-Paradigma gebildete Forschungsgebiet, in dem die Verarbeitung hierarchisch organisierter Stimuli den Gegenstand der Untersuchung darstellt. Abb. 38a zeigt ein Beispiel für einen solchen Stimulus, nämlich den Buchstaben M (globale Ebene), der aus einzelnen Ns (lokale Ebene) zusammengesetzt ist. In bezug auf solche visuellen Eindrücke stellt sich die Frage, ob der Gesamteindruck (das „M“) aus den einzelnen Teilen konstruiert wird oder ob – ganz im Sinn der Gestaltpsychologie – dem Gesamteindruck als Gestalt eine eigenständige und von den Bestandteilen separate Verarbeitung zukommt. Insbesondere auf die Arbeiten von Navon geht die in dieser Hinsicht interessante Hypothese der globalen Präzedenz zurück, nach der der globalen Ebene eine temporale Priorität für die Verarbeitung zukommt.<sup>13</sup>

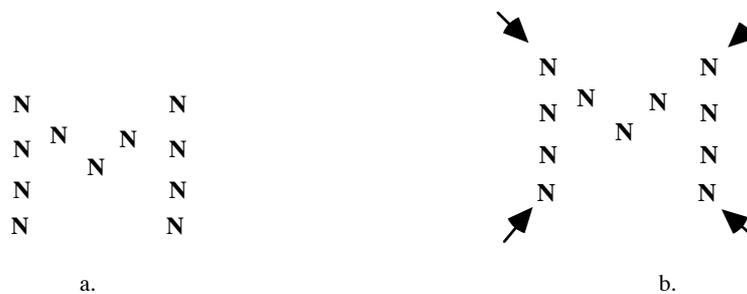


Abb. 38: Hierarchisch organisierte Stimuli im Global/Lokal-Paradigma (Robertson et al. 1993)

Robertson et al. (1993) untersuchten die Rolle von Aufmerksamkeit bei der Identifikation globaler und lokaler Targets in hierarchisch organisierten Stimuli. Ähnlich dem Posner'schen Cuing Paradigma gaben sie den Probanden durch Pfeile einen Hinweis auf

<sup>13</sup> „[P]erceptual processes are temporally organized so that they proceed from global structuring towards more and more fine-grained analysis. In other words, a scene is decomposed rather than built up.“ (Navon 1977:354, zitiert nach Kimchi 1992:26).

bestimmte Orte in einem Muster (s. Abb. 38b), wobei der Hinweis entweder das (jeweils mit einer bestimmten Wahrscheinlichkeit korrekte) Auftreten eines globalen oder eines lokalen Targets vorhersagte. Anhand der Auswirkungen der SOA fanden die Autoren Evidenz für zwei unterschiedliche Mechanismen der Aufmerksamkeit: Einerseits „regionale“ Aufmerksamkeit, die sich in der Verstärkung der gereizten Region und entsprechend zunehmender Erleichterung der Reaktionen äußerte; andererseits „kategorische“ Aufmerksamkeit, die dem Auftreten ebenenspezifischer Effekte zugrunde gelegt wird. Kategorische Aufmerksamkeit drückte sich darin aus, daß, wenn globale Targets erwartet wurden, die Reaktion auf lokale Targets verlangsamt war und umgekehrt. Experimente von Ward<sup>14</sup> zeigen, daß Aufmerksamkeit auf Targets einer Ebene in einem Durchgang n Erleichterung für Targets derselben Ebene in Durchgang n+1 mit sich bringt.

Diese Beobachtungen entsprechen den folgenden Ergebnissen von Paquet/Merikle (1988)<sup>15</sup>: Wenn Versuchspersonen die Aufmerksamkeit auf die globale Information eines hierarchischen Musters an einer bestimmten Stelle im visuellen Feld richteten, so konnte die globale Information eines Musters an einer anderen Stelle nicht ignoriert werden. Wurde umgekehrt die lokale Information eines Musters attendiert, so konnte die lokale Information eines anderen Musters nicht ignoriert werden. Diese Beziehung ist jedoch asymmetrisch: *Globale unattendierte* Objekte werden automatisch einer semantischen Analyse unterzogen, während dies für *lokale unattendierte* Objekte nur dann der Fall ist, wenn die Aufmerksamkeit auf die lokale Ebene eingestellt ist (vgl. hierzu auch Briand 1994).

Kimchi (1992), Robertson et al. (1993) und Kosslyn (1994:98) berichten von mehreren Arbeiten, die diese Effekte auf die Zuordnung unterschiedlicher räumlicher Frequenz (determiniert durch die Anzahl der Objekte in einem festgelegten Bereich) zu den beiden Hierarchieebenen zurückführen. Danach ist niedrige räumliche Frequenz mit der globalen Ebene assoziiert, hohe räumliche Frequenz mit der lokalen Ebene. Diese Analyse ist im Einklang mit ebenenspezifischen Unterschieden der Verarbeitung der beiden Gehirnhälften (sichtbar bei links- bzw. rechthemisphärischen Läsionen, s. Delis/Bihrlé 1989). Jacobs/Kosslyn (1994) mahnen allerdings zur Vorsicht in bezug auf die Interpretation dieser Beobachtungen:

„However, it is worth noting that although the left hemisphere can categorize high spatial frequency gratings faster than the right, and the right can categorize low spatial frequency gratings faster than the left, both hemispheres can *detect* the two sorts of gratings equally well [...]. Such results suggest to us that attentional differences may lie at the heart of these hemispheric effects“ (Jacobs/Kosslyn 1994:382).

Die beobachtete Asymmetrie der beiden Hierarchieebenen in bezug auf die semantische Analyse findet ihre Erklärung in der Arbeit von Boucart et al. (1995), die explizit der Frage nachgehen, welche Objektaspekte (Farbe, Form u.a.) attendiert werden müssen, damit ein Zugriff auf die Identität des Objekts stattfindet. Sie kommen zu dem Schluß, daß „[a]pparently, attention to global information, whatever the physical dimension processed, activates stored representations of objects“ (S. 599).

<sup>14</sup> Berichtet in Robertson et al. (1993).

<sup>15</sup> L. Paquet, P. M. Merikle, „Global precedence in attended and nonattended objects“, *Journal of Experimental Psychology* 14, 89-100.

Kimchi (1992) weist auf ein Problem dieser im Global-/Lokal-Paradigma durchgeführten Analysen hin: Während in der Stimulus-Domäne zwischen zwei Ebenen der hierarchischen Organisation unterschieden wird, weist die perzeptuelle Domäne eine Dreiteilung auf (Ganzform, Teilfiguren, Textur). Demnach ist die Abbildung von der einen in die andere Domäne nicht eindeutig, sondern hängt von der Anzahl und relativen Größe der Teilelemente des Stimulus ab. Dies könnte somit zur Folge haben, daß der Vorteil der globalen Ebene hierarchisch organisierter Muster einen Vorteil der Ganzfigur gegenüber einer Textureinheit reflektiert, und nicht den Vorteil einer globalen gegenüber einer lokalen Eigenschaft.

Die Argumentation Kimchis wird indirekt durch die Arbeit von Palmer/Rock (1994) unterstützt, die die (uneingeschränkte) Priorität von Gestaltbildungsprinzipien bei der perzeptuellen Organisation hinterfragen und somit auch die Hypothese der globalen Präzedenz im Global-/Lokal-Paradigma untergraben. Sie argumentieren gegen eine primäre frühe Anwendung von Gruppierungsprinzipien und plädieren stattdessen für den in Abb. 39 gezeigten Aufbau einer perzeptuellen Organisation. Daran wird ersichtlich, daß Gruppierung der Bildung von sogenannten „Entry-Level-Units“ (nach Figur bzw. Grund unterschiedene zusammenhängende Regionen im visuellen Feld) *nachfolgt*. Entsprechend besteht in bezug auf den Verarbeitungsaufwand ein Unterschied zwischen einer Figur mit einer bestimmten Textur und einem Ganzobjekt, das sich aus einer Anzahl gruppierter Teilelemente zusammensetzt.

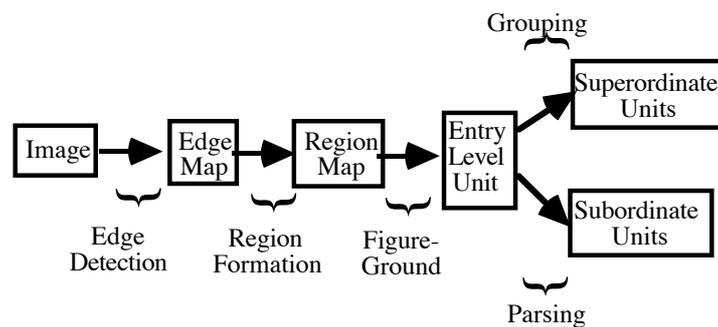


Abb. 39: Entry-Level in der Wahrnehmung (nach Palmer/Rock 1994:42)

#### 4.6 Der operationale Aspekt: Visuelle Suche

Wie bereits erwähnt kommt dem visuell-räumlichen System die für effektives Verhalten unumgängliche Aufgabe der Selektion relevanter Objekte aus dem visuellen Feld zu. Überwiegende Übereinstimmung herrscht in bezug auf dieses Phänomen darüber, daß hierfür zwischen einer präattentiven und einer attentiven Ebene der Verarbeitung unterschieden werden muß. Präattentive Verarbeitung ist durch die folgenden Eigenschaften charak-

terisiert (Theeuwes 1993:99). Erstens ist sie *in der Verarbeitungskapazität unbegrenzt*: In einem visuellen Feld gleicher Distraktor-Elemente wird ein einzelnes Target schnell erkannt (also z. B. ein Kreis mit einem Strich in einem Feld von Kreisen, s. Abb. 40), unabhängig von der Anzahl der Distraktoren (sogenannter „Pop-out“-Effekt). Zweitens operiert präattentive Verarbeitung *räumlich parallel*, da sie gleichzeitig an verschiedenen Stellen des visuellen Feldes stattfindet.<sup>16</sup> Drittens operiert sie *unabhängig von strategischer Kontrolle*, obwohl intentionale top-down Aspekte indirekt einen Einfluß haben können.

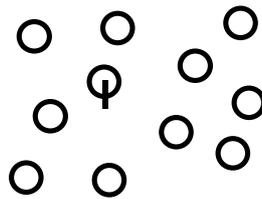


Abb. 40: Das ‚Pop-out‘-Phänomen

Attentive Verarbeitung ist dagegen in der Verarbeitungskapazität begrenzt (führt bei bestimmten Targets mit zunehmender Anzahl der Elemente im visuellen Feld zu einem Anstieg der Reaktionszeit). Bei fehlendem Pop-out des Targets (z. B. das ‚T‘ in Abb. 41a und der Kreis in Abb. 38b) muß daher eine nacheinander ablaufende (serielle) Betrachtung des visuellen Feldes stattfinden, um die entsprechende Auswahl treffen zu können.

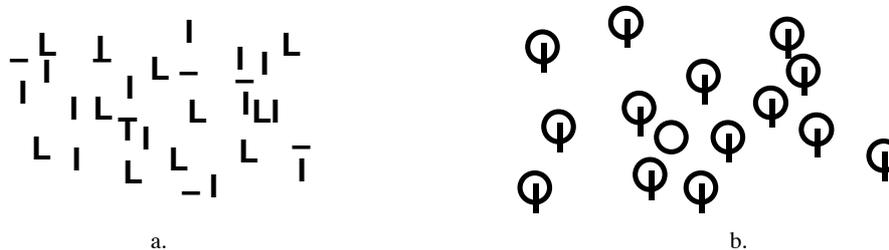


Abb. 41: Beispiele für fehlendes ‚Pop-out‘

Attentive Verarbeitung ist an einen bestimmten räumlichen Ort gebunden (vgl. die Diskussion raum-basierter Aufmerksamkeitsmodelle) und kann außerdem strategisch kontrolliert werden (Ausdehnung und Bewegung des „Spotlights“). Uneinigkeit herrscht allerdings darüber, welche Operationen auf den beiden Ebenen durchgeführt werden (s. o. die „Early“- vs. „Late-Selection“-Debatte).

<sup>16</sup> Theeuwes weist darauf hin, daß „präattentive“ Verarbeitung nicht einheitlich als „vor der bzw. ohne Aufmerksamkeit ablaufend“ interpretiert wird, sondern auch als durch den Zustand verteilter Aufmerksamkeit charakterisiert, im Gegensatz zum Zustand fokussierter Aufmerksamkeit bei attentiver Verarbeitung.

Die „Feature Integration Theorie“ Treismans (FIT, Treisman 1988) bietet eines der gängigsten Modelle visueller Suche. Treisman geht von der vorangegangenen Ebenenunterscheidung und der Annahme aus, daß unterschiedliche Merkmale (Farbe, Orientierung etc.) in verschiedenen Subsystemen verarbeitet werden. Hieraus folgert sie, daß für das Erkennen eines bestimmten Objekts eine räumliche Zuordnung seiner separat und verteilt vorliegenden Eigenschaften – in Abgrenzung zu den Eigenschaften anderer Objekte – vorgenommen werden muß. Sie postuliert daher, daß das Spotlight der Aufmerksamkeit über einer „map of locations“ operiert und den Zugriff auf die mit einem bestimmten Ort assoziierten Merkmale ermöglicht. Dabei werden diejenigen Merkmale betrachtet, die, weil sie sich auf ihren Karten am meisten von konkurrierenden Elementen unterscheiden, am stärksten aktiviert sind (sogenanntes „odd-man-out“-Prinzip, s. Ullman 1984). Diese Merkmale werden attendiert und in eine temporäre Objektrepräsentation integriert (miteinander „verklebt“, s. Abb. 42).

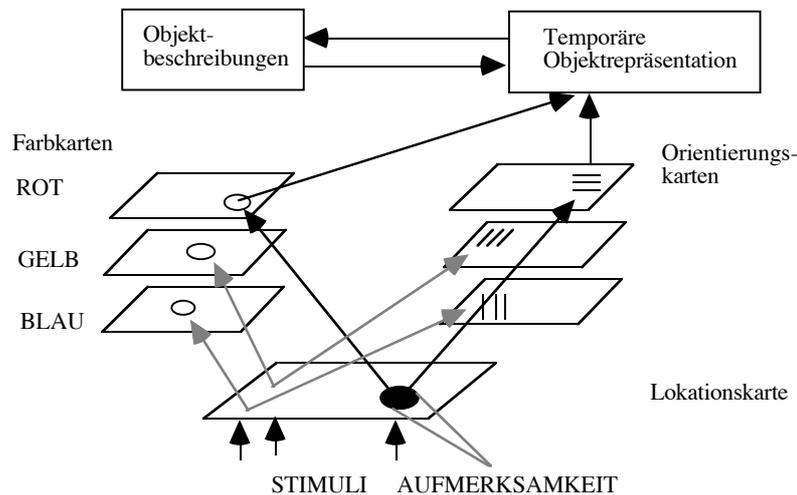


Abb. 42: Das Modell der Feature Integration Theory (nach Treisman 1988:202)

Eine weitere wesentliche, der FIT zugrundeliegende Annahme ist die, daß Ortsinformation von Merkmalen erst durch fokussierte Aufmerksamkeit fixiert wird, daß also Merkmale auf der präattentiven Ebene „frei schweben“. Dies zeigte sich in Experimenten mit verteilter Aufmerksamkeit, in denen Versuchspersonen die Farbe und die Form (Identität) gezeigter Buchstaben falsch zuordneten und so „trügerische Konjunktionen“ („illusory conjunctions“) dieser Aspekte bildeten. Cohen/Ivry (1989) fanden aber, daß illusory conjunctions nur bei *benachbarten*, außerhalb des Spotlights befindlichen Objekten gebildet wurden, so daß zumindest eine grobe Verortung einzelner Merkmale stattgefunden haben mußte. Tsai et al. (1994) und Navon/Ehrlich (1995) stellten ebenfalls einen weniger direkten Zusammenhang zwischen dem Mechanismus fokussierter Aufmerksamkeit und dem Auftreten falscher Merkmalskombinationen fest.

Eine explizite Gegenposition zu diesem Aspekt von FIT nimmt die FINST-Theorie Pylyshyns (1994) ein. Kern dieser Theorie ist die Annahme, daß visuell-räumliche Elemente bereits durch die präattentive Verarbeitung räumlich indexiert sind: „we have gathered converging evidence for the view that there is an early preattentive stage of vision in which the locations of properties in space are selected or *individuated* and *indexed* prior to any serial process being applied to the visual array - in fact even prior to assigning focal attention to places in the array“ (S. 366).

Ein Aspekt der frühen Version der FIT ist die enge Verbindung zwischen der Suche nach konjunktiv definierten Targets und Serialität. Da Merkmalskombinationen erst auf der attentiven Ebene verarbeitet werden, können sie offenbar nicht parallel verarbeitet werden. Entsprechend sollte nach der FIT ein qualitativer Unterschied zwischen der Suche nach einem einzelnen Merkmal (z. B. nach einem „blauen“ Target unter roten und grünen Nicht-Targets) und der Suche nach einer Kombination von Merkmalen (nach einem roten „X“ unter gemischt roten „O“s und grünen „X“en). Zahlreiche Experimente zeigten allerdings, daß Versuchspersonen bestimmte konjunktiv definierte Targets sehr viel schneller in einem visuellen Feld entdecken können als hierdurch vorhergesagt. Aufgrund massiver Gegenevidenz änderten Treisman/Sato (1990) deswegen FIT dahingehend ab (→ „Revised FIT“), daß bei konjunktiver Suche zusätzlich zu dem Mechanismus des Fokus der Aufmerksamkeit ein Merkmalsunterdrückungsmechanismus angenommen wird. Dieser Mechanismus bewirkt, daß nicht-relevante (Distraktor-) Merkmale aktiv unterdrückt werden, so daß, bildlich gesprochen, die Suche auf der „map of locations“ auf die relevanten Objekte eingeschränkt werden kann, was letztendlich eine schnellere Verarbeitung bewirkt.<sup>17</sup>

Duncan/Humphreys (1992) argumentieren prinzipiell gegen eine strikte Abgrenzung paralleler Merkmalsuche und serieller Kombinationsuche. Sie stellten fest, daß die Reaktionszeiten der Probanden von der Unterscheidbarkeit der Targets und Non-Targets abhängig sind: Die Suche ist am einfachsten, wenn die Non-Targets in sich homogen und maximal unähnlich den Targets sind; sie ist schwieriger, wenn die Non-Targets unterschiedlich sind; sie wird zunehmend schwieriger, umso ähnlicher Targets und Non-Targets einander sind. Die Autoren (ebenso wie Wolfe 1994) betonen zusätzlich die Rolle kognitiver (top-down-) Kontrolle der Suche. Duncan/Humphreys nehmen ein sogenanntes Aufmerksamkeitsmuster („attentional template“) an, das eine Beschreibung der gesuchten Information im visuellen Kurzzeitspeicher darstellt. Elemente des visuellen Feldes gewinnen oder verlieren nach ihrer Theorie an Aktivationsstärke (Gewicht), je nachdem ob sie diesem Template entsprechen oder nicht. Außerdem sind Elemente derselben perzeptuellen Gruppe verbunden („weight linkage“), so daß im Falle ähnlicher Non-Targets eine Verbreitung negativer Gewichte und somit deren Unterdrückung stattfindet („spreading suppression“). Diese Vorstellung kombinierter bottom-up und top-down-Aspekte und der Verrechnung der Aktivationsstärke unterschiedlicher Merkmalskarten entspricht den Modellen von Wolfe (1994) und Theeuwes (1993).

<sup>17</sup> Die Suche wird außerdem durch eine flexible Anpassung des Aufmerksamkeitsfensters verbessert, so daß nicht nur einzelne, sondern mehrere Objekte gleichzeitig betrachtet werden können („group scanning hypothesis“).

#### 4.7 Der Repräsentationsaspekt: Object-Files

Es ist bereits erwähnt worden, daß Repräsentationsaspekte in der hier betrachteten Aufmerksamkeitsforschung eine eher marginale Rolle spielen. Dies wird daran deutlich, daß in den vorangegangenen Abschnitten vorwiegend Untersuchungen zu Mechanismen der Aufmerksamkeit bzw. zu entsprechenden Verarbeitungsaspekten vorgestellt worden sind. Es existieren allerdings bestimmte Bereiche, in denen der Repräsentationsaspekt visueller Aufmerksamkeit explizit diskutiert wird. Hierzu gehört die Frage, wodurch die anscheinende Leichtigkeit zu erklären ist, mit der wir trotz einer objektiv dynamischen Wahrnehmung der Welt (durch Bewegung der Objekte in der Welt, durch Eigenbewegung, oder durch unterschiedliche Informationsselektion während aufeinanderfolgender Fixationen) einen subjektiv stabilen Eindruck erhalten, der sich in dem Erkennen der „Statik“ unbeweglicher Objekte und der „Kontinuität“ sich bewegender Objekte äußert.

Zur Untersuchung der Objektcontinuität existiert das sogenannte „Vorschau-Paradigma“ („preview paradigm“), in dem den Probanden vor der Präsentation des Targets ein Display gezeigt wird, das das Target enthält oder nicht. Hierbei lassen sich Vorteile solcher previews beobachten, die auf mindestens zwei unterschiedliche Arten erklärt werden können (s. Henderson/Anes 1994). Einerseits könnten durch den preview entsprechende Objekttyp-Repräsentationen im Langzeitgedächtnis aktiviert werden, so daß preview-Vorteile als auf allgemeinen Priming-Phänomenen beruhend charakterisierbar sind (Pollatsek et al. 1990). Andererseits könnten diese Vorteile darauf zurückzuführen sein, daß bei einem Preview des Targets eine temporäre Repräsentation gebildet wird, die der Integration von Information aufeinanderfolgender Sichten dient.

Die Theorie der object files von Kahneman und Treisman (Kahneman/Treisman 1992) ist dem zweiten Erklärungstyp zuzuordnen. Sie basiert auf der grundlegenden Unterscheidung von „wahrnehmen“ und „erkennen“: Wir können einerseits Objekte wahrnehmen, ohne zu erkennen von welchem Typ diese Objekte sind; wir können andererseits ein wahrgenommenes Objekt aufgrund von Merkmalsveränderungen unterschiedlichen Typen zuordnen, ohne dabei die Kontinuität des Objekts selbst infrage zu stellen (Morphing-Beispiele; das Superman-Beispiel, s.o.). Sie betrifft außerdem das Phänomen anscheinender Bewegung (apparent motion): unter bestimmten zeitlich-räumlich beschränkten Bedingungen nehmen wir ein Objekt als sich von A nach B bewegend wahr, auch wenn nur zwei „snapshots“ des Objekts am Ort A und am Ort B gesehen werden. Hierbei ist entscheidend, daß dieser Eindruck erst *nach* der Darbietung des Objekts am Ort B entsteht.

Zur Erklärung dieser Phänomene nehmen die Autoren an, daß das visuelle Feld in eine Menge perzeptueller Objekte (und einen undifferenzierten Hintergrund) geparkt wird und daß die Verarbeitung dieser Objekte zur Erstellung episodischer Tokenrepräsentationen (sog. *object files*) führt, die primär durch ihre räumlichen und zeitlichen Charakteristika adressiert und somit variabel bzgl. anderer Merkmale sind. Drei Prozesse bewirken nach dieser Theorie die beschriebenen Phänomene: Ein Korrespondenzprozess, der bestimmt, ob ein bestimmtes Objekt „neu“ ist oder einem vorher an einem anderen Ort wahrgenommenen entspricht; ein Nachschau- („reviewing“) Prozess, der die Charakteristika des unmittelbar vorhergehenden Objektes bereitstellt; ein Vervollständigungs- („impletion“) Prozess, der die aktuelle und vorhergehende Information so miteinander verrechnet, daß der Ein-

druck eines Ortswechsel oder anscheinender Bewegung entsteht. Eine wesentliche Rolle beim reviewing spielt die Zuwendung von Aufmerksamkeit auf das Target. Hierdurch wird einer der aktiven object files ausgewählt, was im Falle eines Matches mit dem Target zum beobachteten preview-Vorteil und im Falle eines Mismatches zu Interferenzen führt.

Den Kern der Theorie Kahneman/Treismans bildet das Ergebnis, daß die räumlich-zeitliche Kontinuität ausschlaggebend für die Erleichterung des Erkennes des Targets ist: Diese trat in den Experimenten z.B. dann ein, wenn das Target (ein Buchstabe) sowohl im preview-Display als auch im Target-Display an dasselbe perzeptuelle Objekt gebunden war (sich im selben Kästchen befand), im Gegensatz zu dessen Erscheinen an einer anderen Stelle im Target-Display. Diesen lokationsspezifischen Effekt konnten Pollatsek et al. (1990) gerade nicht finden, so daß sie die preview-Vorteile auf eine Aktivierung sogenannter „Objekt-Detektoren“ (abstrakten Objekttyp-Repräsentationen) zurückführten. In einem experimentellen Vergleich der Object-File-Theorie mit der Detektor-Priming-Theorie kommt Henderson (1994) allerdings zu dem Ergebnis, daß eine Synthese der beiden Erklärungsmodelle eine plausible Alternative darstellt, in der die object files neben der direkten Aktivierung abstrakter Typinformation eine zentrale Rolle bei der Objekterkennung spielen.

Die Annahme von object files wird unterstützt durch Untersuchungen zur „*repetition blindness*“ (RB, s. Park/Kanwisher 1994) und zum „*attentional blink*“ (AB, Chun 1997). RB bezeichnet den Effekt, daß Versuchspersonen bei schnell aufeinanderfolgender Präsentation zweier gleicher Stimuli (innerhalb des rapid serial visual presentation Paradigmas) das Vorkommen des zweiten Stimulus häufig nicht bewußt wird. Zur Erklärung dieses Phänomens nehmen Kanwisher/Park an, daß zwar ein „Erkennen“ des zweiten Stimulus (Zuordnung visuellen Inputs zu abstrakten Repräsentationen), aber keine „Individuierung“ (Bindung der aktivierten Typinformation an ein neues „token“ oder object file) stattfindet.<sup>18</sup> Komplementär dazu bezeichnet AB den Effekt, daß das Berichten des Auftretens zweier unterschiedlicher Targets defizitär ist, wenn beide in einem Zeitfenster von 500ms auftreten. Dies läßt sich nach Chun (1997) dadurch erklären, daß die *Bildung* der für ein Berichten vorauszusetzenden object files Ressourcen in Anspruch nimmt, deren Wiederherstellung eine gewisse Zeit (refractory period) dauert und so die Bildung eines zweiten object files beeinträchtigt.

Park/Kanwisher stellen vor diesem Hintergrund bzgl. der Rolle von Aufmerksamkeit eine interessante Hypothese auf, die als „Selection-for-report“ bezeichnet werden kann:

„Recognition itself is not capacity-limited, but the transfer of recognized items to a subsequent buffer necessary for report is a serial, attention-requiring process“ (S. 516).

<sup>18</sup> Allerdings bestehen Unterschiede zwischen den Theorien Kahneman/Treismans und Park/Kanwisher in bezug auf die Eigenschaften bzw. die Funktion von object files/tokens. Kahneman/Treisman betrachten object files als eine integrative Instanz („Whenever a change in visual input is detected, current information about changing or reappearing objects must be assigned to existing object files; if this fails, a new file must be set up.“, S. 179). Park/ Kanwisher weisen explizit auf die Funktion eines Tokens als Individuierungskonstrukt eines Stimulus hin und erklären RB so als Bindungsproblem existierender Token zu Types: „[...] subjects evidently knew something was there; they just did not know what it was. This is consistent with the existence of a token but the lack of a type-to-token link for the repetition-suppressed item“ (S. 516). Demnach sind die beiden Konstrukte nicht äquivalent. Stattdessen scheint eine größere Nähe des ‚Token‘-Begriffs zu dem des ‚FINST‘ in dem Ansatz Pylyshyns vorzuliegen, der gerade durch die Einschränkung auf Individuiertheit und räumliche Adressiertheit eines visuellen Objekts ausgezeichnet ist.

#### 4.8 Mechanismen der Aufmerksamkeit und Raumrepräsentation: Der neuropsychologische Aspekt

„[The human hippocampus may] house the human cognitive map in the sense of storing the information from which it is constructed but the clinical evidence suggests that the neural equivalents of surveyor, cartographer and navigator are located in the right parietal lobe.“ (Ratcliff 1987:85)

Es besteht mittlerweile ein breiter Konsens darüber, daß ausgehend vom primären visuellen Cortex (Area V1) neuronale Aktivierung entlang zweier unterschiedlicher Pfade weitergeleitet wird: Entlang der dorsalen Route, die im posterioren parietalen Cortex endet, und entlang der ventralen Route, die im inferotemporalen Cortex endet. Insbesondere anhand von Läsionsstudien konnte gezeigt werden (Ungerleider/Mishkin 1982), daß der dorsale Pfad mit der Verarbeitung räumlicher Funktionen („where“-Aspekt) und der ventrale Pfad mit der Verarbeitung von Objekteigenschaften („what“-Aspekt) assoziiert werden kann. Bzgl. des parietalen Cortex weist Paillard (1987) darauf hin, daß hier objekt-relative („allozentrische“) räumliche Beziehungen repräsentiert werden, die einerseits von in retinotopischen Koordinaten organisierten visuellen Repräsentationen des visuellen Cortex, und andererseits von Betrachter-relativen („egozentrischen“) Repräsentationen im frontalen Bereich abgegrenzt werden können. Einer der klassischen, für parietale Schädigungen typischen Fälle visueller Disorientierung (auch „Bálints Syndrom“ genannt) wird von Holmes (zitiert in Ratcliff 1982:305)<sup>19</sup> geschildert, dessen Patienten beidseitige, durch Projektile hervorgerufene, Verletzungen der hinteren Hirnregionen aufwiesen und deren offensichtlichstes Symptom „the inability to determine the position in space, in relation to themselves, of objects which they saw distinctly“ war, obwohl keine Beeinträchtigung der Fähigkeit zur Bewegung der Augen selbst vorlag:

„When I held up a knife in front of one man he said once 'That's a pocket-knife' but though his eyes were directed on it he stretched out his arm in a totally wrong direction when he was told to take hold of it. Another man struck my face with his hand as he attempted to point to a pencil which I held two or three feet to my right side“

Außer dieser Art visueller Mislokalisierung können sich parietale Schädigungen außerdem in konstruktiver Apraxie (Unfähigkeit, räumliche Relationen in Form von Modellen oder durch Zeichnungen wiederzugeben), visueller Disorientierung (Unfähigkeit, relative Größen, Längen, oder Positionen von Objekten anzugeben), Verlust räumlichen Gedächtnisses (z.B. mit Beeinträchtigung der Fähigkeit, Routen zu finden) und *Aufmerksamkeitsdefiziten* ausdrücken (s. Andersen 1987). Neben verbesserten Techniken des „brain imaging“ (s. Posner/Raichle 1994, Mangun et al. 1993), d. h. Verfahren zur Visualisierung der Gehirnaktivität, ist es insbesondere die Neglect-Forschung (Bisiach 1993), die zu einem Fortschritt in der Erforschung von Aufmerksamkeitsphänomenen sowie dem Verhältnis von Aufmerksamkeit und Raumrepräsentation geführt hat.

<sup>19</sup> G. Holmes, „Disturbances of visual space perception“, *British Medical Journal*2 (1919), 230-233.

#### 4.8.1 Raum-basierter Neglekt

Seit mehr als einem Jahrhundert existieren Untersuchungen über Personen mit einer Schädigung des parietalen Bereichs, die als „Unilateraler Neglekt“ bezeichnet wird. Solche Personen lassen Information über diejenige Seite des Raumes unberücksichtigt, die der Läsion gegenüberliegt – üblicherweise der rechten Gehirnhälfte („full-fledged neglect is usually the consequence of a lesion located in the right hemisphere“, Bisiach/Vallar 1988: 195). Dieser Umstand wird daran sichtbar, daß sie bei vorgegebenen oder memorierten räumlichen Gegebenheiten nur die rechte Seite des Raumes wiedergeben, sei es durch Zeichnungen oder durch verbale Beschreibungen.<sup>20</sup>

Entscheidend für den Zusammenhang von parietaler Schädigung, visuell-räumlicher Aufmerksamkeit und Raumrepräsentation ist die Tatsache, daß bei diesen Fällen anscheinender Beeinträchtigung räumlicher Repräsentation tatsächlich eine Störung der Aufmerksamkeit vorliegt („Unilateral neglect stands apart from the other spatial disorders in that the neglecting patient is not simply unable to integrate stimuli falling into one half of space into a coherent spatial framework but appears to be unaware of their existence“, Ratcliff 1982:322). Dies wird durch Untersuchungen belegt (Goldberg 1982), in denen Einzelzelleitungen parietaler Neuronen im Gehirn von Affen vorgenommen wurden. Von solchen Neuronen konnte zunächst gezeigt werden, daß sie mit irgendeiner Reaktion des Tieres auf einen Stimulus in Verbindung stehen. Zusätzlich ergaben die Vergleiche elektrophysiologischer Aktivitäten von einfacher Fixation, von Aufmerksamkeitszuwendung auf einen entsprechenden peripheren Stimulus und von einer Augenbewegung zu diesem Stimulus, daß die jeweiligen Neuronen genau dann verstärkt feuerten, wenn ausschließlich Aufmerksamkeitszuwendung, aber keine Augenbewegung vorlag: „In the posterior parietal cortex this process of enhancement occurs whenever the animal attends to the stimulus, not merely when it makes an eye movement to it, and is therefore useful for the processing underlying selective visuospatial attention“ (ibid., S. 297).

Diese Ergebnisse entsprechen offenbar den von Posner untersuchten Zuwendungen visuell-räumlicher Aufmerksamkeit. Der parietale Cortex ist somit nicht in einfacher Weise als ein für die allozentrische Repräsentation von Raum funktionaler Bereich anzusehen, sondern kodiert außerdem einen spezifischen attentionalen Mechanismus. Räumlicher unilateraler Neglekt, der auf einer Schädigung parietaler Bereiche beruht, kann nach Allport wie folgt charakterisiert werden: „[It] is not restricted to a fixed region of egocentric space (e.g., the contralesional visual hemifield), but shows a left-to-right, continuous *gradient* or bias of attentional priority [...] In other words the disturbance appears to reflect an enhanced attentional priority, and stronger attentional engagement, in one direction, as much as a diminished priority towards the other“ (Allport 1993:196f).

Auffällig ist die Asymmetrie der als attentionale Defizite enttarnten Neglekt-Phänomene bezüglich der beiden Gehirnhälften. Sie treten vorwiegend bei Verletzungen der rechten (allgemein als der auf die Verarbeitung räumlicher Information spezialisiert angesehenen) Hälfte auf (Bisiach/Vallar 1988:198). Ládavas et al. (1994) vergleichen verschiedene Erklärungshypothesen für diese Asymmetrie. Nach einem Modell verfügt jede Gehirnhälfte

<sup>20</sup> Hierbei handelt es sich um Neglekt des extrapersonalen Raumes. Es sind außerdem Fälle bekannt, in denen sich die Nichtbeachtung auf den intrapersonalen Raum (z. B. auf eine entsprechende Körperhälfte) bezieht (Bisiach/Vallar 1988:196).

über eigene Aktivierungssysteme, die Zuwendungsreaktionen zum jeweils contralateralen Halbraum auslösen. Die Asymmetrie wird hier dadurch erklärt, daß die linke Hemisphäre nur Zuwendungen zur rechten kontrollieren kann, während die rechte Hemisphäre Zuwendungen auf Stimuli in beiden Halbräumen bewirken kann. Nach einem anderen Modell ist jede Gehirnhälfte für den Aufmerksamkeitswechsel in die Richtung der jeweils anderen Hälfte zuständig, wobei die rechte Gehirnhälfte schwächer in bezug auf die Generierung solcher Wechsel ist als die linke (daher eine Unterdrückung der linksgerichteten Aufmerksamkeitswechsel bei rechtshemisphärischen Läsionen). In eigenen Untersuchungen mit split-brain-Patienten (Patienten mit durchtrenntem Corpus Callosum), kommen die Autoren zu Ergebnissen, die eher das zweite Modell stützen. Sie fanden allerdings keine Unterschiede in der Stärke der Zuwendungsrichtungen von Aufmerksamkeit. Stattdessen ergab sich, daß die rechte Gehirnhälfte über einen relativ höheren Wachsamkeitszustand verfügt.

In Experimenten, in denen Corbetta und Kollegen anhand von PET scans die Gehirnaktivierung bei Aufmerksamkeitswechseln untersuchten (berichtet in Posner/ Raichle 1994:160f), wurde allerdings das erste Modell klar bestätigt: Beide Gehirnhälften (genauer: die parietalen Bereiche) waren aktiv, wenn Aufmerksamkeitswechsel im rechten visuellen Halbfeld stattfanden; hingegen war nur der rechtshemisphärische parietale Bereich aktiv, wenn entsprechende Wechsel im linken visuellen Halbfeld stattfanden. Insbesondere: „It did not seem to matter whether subjects were shifting [attention] in a rightward or leftward direction“ (Posner/Raichle 1994:161).

Obwohl die attentionale Verarbeitung räumlicher Repräsentationen bei parietalen Neglekt-Patienten also in spezifischer Weise gestört ist, bedeutet dies nicht unbedingt eine Beeinträchtigung der vorliegenden Raumrepräsentation. Bisiach/Luzatti (berichtet in Andersen 1987:485) baten solche Patienten, aus dem Gedächtnis bestimmte Landmarken zu beschreiben, die direkt am Piazza del Duomo in Mailand lagen. Hierzu sollten sie zunächst eine bestimmte Perspektive einnehmen (Blick auf den Eingang der Kathedrale). Wie erwartet, beschrieben die Patienten nur diejenigen Landmarken, die sich auf derselben Seite (ipsilateral) der vorliegenden Läsion befanden. Danach wurden sie gebeten, eine entgegengesetzte Perspektive (von den Stufen der Kathedrale aus) einzunehmen. Wieder beschrieben sie nur eine Seite des Platzes, dabei erwähnten sie aber Objekte, die sie vorher ignoriert hatten, und sie ignorierten Objekte, die sie vorher genannt hatten. Während die Memorierung bei diesen Personen vor den Gehirnverletzungen stattgefunden hatte, berichten Guariglia et al. (1993) von einem Patienten – allerdings mit rechts-frontaler Läsion –, der keine allgemeinen Beeinträchtigungen von Wahrnehmung oder Imagination aufwies. Dieser zeigte ebenfalls den contralateralen Neglekt von Landmarken an bekannten Plätzen. Zusätzlich zeigte er dasselbe Phänomen aber auch bzgl. der in einem ihm unbekanntem Raum befindlichen Objekte: Nachdem er die Konstellation memoriert hatte, beschrieb er aus dem Gedächtnis jeweils nur die eine Seite des Raumes.

#### 4.8.2 Die Verarbeitung hierarchischer Stimuli

Eine weitere frappierende Asymmetrie der beiden Gehirnhälften läßt sich bei der Wiedergabe hierarchisch organisierter Stimuli beobachten (s. hierzu Robertson/Lamb 1991, Delis/Bihrlé 1989). Während Patienten mit linkshemisphärischen Läsionen in der Regel

nur Aspekte der globalen Ebene (Gestalt eines Objekts) zeichnerisch wiedergeben können, sind die Zeichnungen rechtshemisphärischer Patienten meistens in bezug auf die Gestalt der Figur verzerrt, weisen dafür aber die Aspekte der lokalen Ebene (die Teile der Figur) auf (s. o., Abb. 38). Robertson/Lamb zeigen, daß in beiden Hemisphären sowohl perzeptuelle als auch attentionale Faktoren bei der Verarbeitung hierarchischer Stimuli beteiligt sind und daß diese Stimuli gleichzeitig auf beiden Ebenen (global, lokal) verarbeitet werden. Die Aufmerksamkeitsprozesse, die nach Ansicht der Autoren für die Zuweisung von Ressourcen auf die ebenenspezifischen Mechanismen verantwortlich sind (vgl. die „kategorische“ Aufmerksamkeit bei Robertson et al. 1993), lokalisieren sie im parietalen Bereich, eine Annahme, die von Delis/Bihle bestätigt wird: „The present study suggests that the inferior parietal region is critical for the controlled distribution of attention over the visual field and between hierarchical levels.“ (S. 29)

#### 4.8.3 Neglekt in unterschiedlichen Referenzsystemen

Eine der erstaunlichsten Entdeckungen der Neglekt-Forschung ist die, daß unilateraler Neglekt nicht auf eine Vernachlässigung eines visuellen Halbfeldes beschränkt ist, sondern auch in bezug auf interne Repräsentationen auftritt („representational neglect“, s. Bisiach/Vallar 1988:211f). Bisiach und Kollegen (berichtet in Bisiach 1993:438f) wiesen dies in einem Experiment nach, in dem Patienten mit Links-Neglekt „gleich/unterschiedlich“-Bewertungen bzgl. Paaren bedeutungsloser Objekte abgeben mußten, wobei die Form der Objekte identisch war oder aber sich die Unterschiede entweder auf der linken oder rechten Seite befanden. Die Entdeckung linksseitiger Unterschiede zeigte sich insgesamt als beeinträchtigt, und zwar insbesondere auch in einer Bedingung, in der die Objekte jeweils nur partiell sichtbar waren. Partielle Sichtbarkeit wurde dadurch hergestellt, daß die Objekte hinter einer Wand vorbeibewegt wurden und nur durch einen vertikalen Schlitz ausschnittsweise betrachtet werden konnten. Der auftretende Neglekt bezog sich dementsprechend nicht auf eine perzeptuelle, sondern visuell-räumliche Repräsentation.

Allport gibt zwei weitere Beispiele: Ein Patient mit linksseitigem Neglekt könne Stimuli auf der bzgl. der *Gravitationsachse* definierten linken Seite mißachten, auch wenn der Kopf horizontal gekippt ist; andererseits könne ein solcher Neglekt auch in bezug auf die *Hauptachse* eines gekippten Objekts auftreten, trotz aufrechter Haltung (Kopfposition) des Betrachters. Den wohl eindrucksvollsten Beleg liefert jedoch ein Fall von räumlichem Neglekt in einem orthographisch definierten Referenzsystem:

„The patient studied in detail by Caramazza and Hillis failed to read the *terminal* (i.e., in canonical, alphabetic representation, the 'right') half of words, regardless of whether the word was presented visually in normal left-to-right orientation, or was mirror-reversed, or even if the words were orally spelled to the patient. Thus, hemineglect was manifested within what appears to be a word-centered, orthographic space, which is evidently *not* retinotopic.“ (Allport 1993:198)

Einen ähnlichen Fall für Neglekt auf einer *spezifischen* Repräsentationsebene schildern Marshall/Halligan (1995). Ihre Patientin konnte zwar die globale Figur eines hierarchischen Stimulus verbal und zeichnerisch wiedergeben, sie konnte jedoch nur die lokalen Elemente auf der rechten Seite der Figur „ausstreichen“ (auch wenn sie vorher die Gesamtfigur korrekt gezeichnet hatte): „J.R. can perceive the whole forest but cannot use that

percept to search for and cut down the trees on the left thereof“ (S. 523). Dies ist offenbar ein Beispiel dafür, daß scheinbar vorhandene Information nicht verwendet werden kann. Interessanterweise kann die konverse Konstellation (scheinbar nicht vorhandene Information kann verwendet werden) ebenfalls eintreten. Vallar et al. (1994) erwähnen den Fall einer Patientin, die zwei Zeichnungen eines Hauses als gleich beurteilte, obwohl das eine Haus auf der linken Seite Flammen (eingezeichnet) hatte. Wurde sie jedoch gefragt, in welchem Haus sie am liebsten wohnen würde, so wählte sie konsistent das nicht-brennende Haus. „This suggests that, at least in some neglect patients [...], overtly neglected left-sided information may undergo processing deep enough to allow the extraction of structural or even semantic features.“ (S. 377; vgl. hierzu auch Farah 1990).

#### 4.8.4 Objekt-basierter Neglekt

Analog zur Unterscheidung raum-basierter und objekt-basierter Aufmerksamkeit gibt es neben der Evidenz für raum-basiertem Neglekt auch Hinweise auf objekt-basierten Neglekt (Behrmann/Tipper 1994, Humphreys et al. 1994). Z. B. ist es für Patienten mit dem sogenannten „Balint’s Syndrom“ charakteristisch, daß sie nur das im Zentrum der Aufmerksamkeit befindliche Objekt „sehen“ können. Humphreys/Riddoch (1993: 159) berichten von einem Patienten Lurias, der nur ein einziges Dreieck sah, wenn zwei Dreiecke unterschiedlicher Farbe übereinandergelegt waren. Waren die Dreiecke jedoch von gleicher Farbe, „sah“ er einen Davidsstern. Ein anderer Patient hatte keine Schwierigkeiten, Muster mit sich überlappenden Objekten (wie die Olympischen Ringe) zu kopieren, im Gegensatz zu Mustern mit separaten Objekten (obwohl diese den gleichen Raum einnahmen).

In ihrem Experiment präsentierten Humphreys/Riddoch zwei Patienten mit Balint’s Syndrom Displays mit 32 farbigen Kreisen (entweder alle rot oder grün, oder zur Hälfte rot und zur anderen grün) und schwarzen Balken mit der Aufgabe zu entscheiden, ob Kreise unterschiedlicher Farbe vorhanden waren. Drei Bedingungen wurden unterschieden: *gemischte* Bedingung (verschiedenfarbige Kreise sind durch schwarze Balken verbunden), *einzelne* Bedingung (gleichfarbige Kreise sind durch Balken verbunden) und *zufällige* Bedingung (Kreise sind nicht verbunden). Es zeigte sich, daß die Performanz bei beiden Patienten in der gemischten Bedingung am besten war, und zwar mit signifikantem Abstand zu den etwa gleichen anderen Bedingungen. Mit anderen Worten, erst die Verbindung zweier Kreise zu einem perzeptuellen Objekt (vgl. das Prinzip der „uniform connectedness“ bei Palmer/Rock 1994) bewirkte die Befindlichkeit zweier Kreise im Bereich der Aufmerksamkeit und die nachfolgende Entdeckung von Kreisen unterschiedlicher Farbe.

Ein spezifische, auf die bilaterale Schädigung parietaler Regionen zurückführbares, sehr ähnliches Krankheitsbild ist die „dorsale Simultanagnosie“. Entsprechende Patienten verfügen zwar über ein intaktes visuelles Feld, sie sind jedoch nicht in der Lage, andere Objekte als das gerade attendierte zu sehen, d.h. einen Aufmerksamkeitswechsel in *irgendeine* Richtung durchzuführen. Allport erwähnt den Fall eines Patienten, der beim Anzünden einer Zigarette die wenige Zentimeter entfernte angebotene Flamme nicht wahrnehmen konnte, da seine Augen auf die Zigarette fixiert waren (Allport 1993:199).

#### 4.8.5 Mechanismen der Aufmerksamkeit und ihre Lokalisierung

Posner/Raichle (1994) präsentieren einen elaborierten Vorschlag dazu, durch welche Mechanismen die beschriebenen Aufmerksamkeitsphänomene erklärt werden können und durch welche neuronalen Systeme die jeweils entsprechende Funktionalität implementiert ist. Sie unterscheiden drei Aufmerksamkeitsnetzwerke: Ein Netzwerk zur visuellen Orientierung (Aufmerksamkeitszuwendung), ein ausführendes (exekutives) Netzwerk, und ein Wachsamkeitsnetzwerk (vigilance network). Das Netzwerk der Aufmerksamkeitszuwendung wird durch das Zusammenspiel der drei Operationen DISENGAGE, MOVE und ENGAGE gebildet. Das Funktionieren visuell-räumlicher Aufmerksamkeit besteht danach in einem sich stetig wiederholenden Lösen der Aufmerksamkeitszuwendung von einem Ort (DISENGAGE), dem Bewegen (des attentionalen Spotlights) zu einem anderen Ort (MOVE) und der Zuwendung an diesem Ort (ENGAGE).

Insbesondere vor dem Hintergrund der beschriebenen Neglekt-Phänomene schlagen die Autoren als Ort für die DISENGAGE-Operation den parietalen Cortex vor. Konvergierende Evidenz weist darauf hin, daß Schädigungen in diesem Bereich zu einer Beeinträchtigung contralateraler Orientierung führen. Anhand von normalen Reaktionen auf ein Target in beiden Halbfeldern konnte allerdings gezeigt werden, daß die Aufmerksamkeitszuwendung selbst nicht unbedingt gestört war. Wird zudem ein Cue vorgegeben, so erhöht sich die Reaktionszeit auf das Target erst dann, wenn die Aufmerksamkeit von der Cue-Position in contraläsionaler Richtung zum Target gewandt werden muß (dabei ist es irrelevant, in welchem Halbfeld sich Cue und Target befinden) – eine Erscheinung, die von Posner/Raichle als Mißlingen des Loslösens von der Cueposition gedeutet wird. Als Ort für die MOVE- und die ENGAGE-Operation schlagen die Autoren jeweils entsprechend den Superior Colliculus im Mittelhirn und die Pulvinar im Thalamus vor.

Das ausführende Netzwerk hat nach Posner/Raichle die Aufgabe, das Bewußtwerden über das Vorliegen eines Objekts zu bewirken und somit dessen Identifizierung zu ermöglichen. Hierbei werden auch top-down-Aspekte, d.h., ob das Objekt einer bestimmten Zielvorgabe entspricht, relevant. Insbesondere anhand von PET-Studien läßt sich nachweisen, daß bei aktiver Einwirkung auf einen wahrgenommenen Reiz der Bereich des anterior cingulate gyrus aktiviert wird. Dieser Bereich ist verbunden mit den vorderen Hirnregionen, denen die Funktionalität des Arbeitsgedächtnisses (sowohl räumliches als auch verbales) zugeschrieben wird. Aus den Effekten von Läsionen dieser Region läßt sich schließen, daß dort attentionale Kontrolle implementiert ist, die sowohl das Arbeitsgedächtnis als auch die visuelle Verarbeitung beeinflussen kann.

Wiederum auf der Basis von PET Scans kommen die Autoren zu dem Schluß, daß im rechten vorderen und parietalen Bereich das Wachsamkeitsnetzwerk lokalisiert ist, das komplementär zum exekutiven Netzwerk operiert. Wenn sich also die Aktivität im Bereich des anterior cingulate gyrus verringert, erhöht sich entsprechend die Aktivität in den angesprochenen Regionen der rechten Gehirnhälfte. Dies dient dazu, eintreffende Reize besser zu verarbeiten, um möglichst schnell auf sie reagieren zu können.

Humphreys/Riddoch (1994) bieten ein abstrakteres Modell des attentionalen Netzwerks visueller Orientierung an (Abb. 43), demgemäß die drei abgebildeten Mechanismen in jeweils gegenseitig hemmender Beziehung zueinander stehen. Die Operationen Posner/Raichles ergeben sich nach diesem Modell als graduell ineinander übergehende Zustände

des Netzwerks. Außerdem ist eine feinere Analyse der Phänomene bei parietaler Läsion möglich: „Disengagement“ kann sich einerseits als Defizit des Orientierungsmechanismus (keine „Attraktion“ durch den contralateralen Stimulus möglich) oder durch übermäßige Inhibition durch den Mechanismus des Aufrechterhaltens der Aufmerksamkeit (abnormales „Verhalten“ der Aufmerksamkeit an der ipsilateralen Position) ergeben. Zur Unterscheidung raumbasierter Orientierens und objektbasierter Selektion schlagen die Autoren die in Abb. 44 dargestellte schematische Beziehung eines Systems zur Objekterkennung und des Aufmerksamkeitssystems vor. Nach diesem Modell ergibt sich visuelle Orientierung aus einem stabilen Zustand der Beziehung. Tritt hingegen verstärkt Feedback vom Objekterkennungssystem auf, so führt dies zu objektbasierter Selektion.

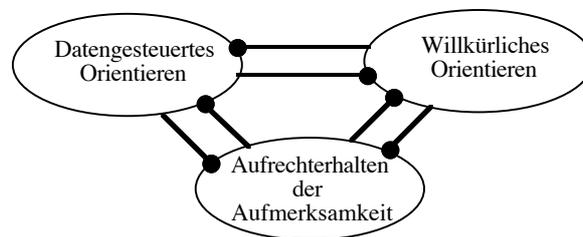


Abb. 43: Attentionales Netzwerk sich gegenseitig hemmender Mechanismen (Humphreys/Riddoch)

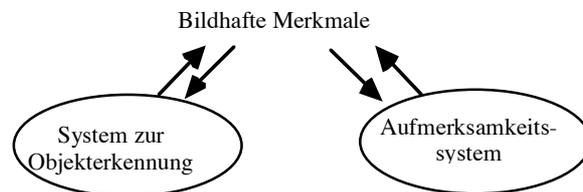


Abb. 44: Modellierung raum- und objektbasierter Selektion durch variierenden Einfluß zweier Systeme (Humphreys/Riddoch)

#### 4.9 Zusammenfassung

Die hier vorgestellten Untersuchungen und Experimente stellen Teile eines Puzzles dar, das bislang noch kein Gesamtbild einer Theorie der selektiven fokussierten visuell-räumlichen Aufmerksamkeit (im folgenden: SFA) ergibt. Trotzdem lassen sich einige wesentliche Aussagen ableiten, die im Rahmen der vorliegenden Arbeit relevant sind, da sie als Grundlage

für die weiteren Betrachtungen räumlicher Relationen im Bereich höherer kognitiver Verarbeitung dienen können.

Es zeigt sich zunächst, daß SFA eine zentrale Rolle in dem Zusammenspiel von Wahrnehmung, Repräsentation und Handeln spielt: Ihre Selektivität ermöglicht Beschränkungen im Informationsaustausch von Teilsystemen („selection for action“, „selection for report“), die die Effizienz der Verarbeitung gewährleisten und die Effektivität des Gesamtsystems sichern. SFA ist dabei ein eigenständiges Phänomen, das sowohl in seiner abstrakten Struktur untersucht als auch an spezifischen neuronalen Korrelaten festgemacht werden kann. Dies impliziert jedoch nicht die Annahme eines lokalisierbaren kausalen Homunkulus (oder eine „cause theory“ im Sinne Johnston/Darks) und somit einer Instanz, die über Repräsentationen operiert, sondern ist, wie Bisiach vorschlägt, möglicherweise integraler Bestandteil von ‚Repräsentation‘: „[...] I claim that attention is a concept rooted in, and abstracted from, the intrinsic dynamics of representation and not a process acting upon it“ (Bisiach 1993:456).

Die Zuwendung visueller Aufmerksamkeit scheint auf noch nicht identifizierte saliente Merkmalscluster in/auf einem räumlichen Medium (der „map of locations“) zu erfolgen („shape integration takes place within an attended region of space at the boundaries of an object [...] Once the attention is set to an object's boundary, the internal features of the object are integrated, and recognition processes begin“, Boucart/Humphreys 1992:804). Dies können einzelne visuelle Elemente („Blobs“), aber auch Gruppen von Blobs oder markante Merkmalsunterschiede (Begrenzungen von Objekten) sein. Eine solche *raum-basierte Aufmerksamkeit* führt zur Etablierung räumlich-zeitlich adressierter „object files“, deren phänomenale Inhalte (Merkmale) sowie kategoriale Zuordnung variieren können. Davon zu unterscheiden ist die *objektbasierte Aufmerksamkeit*, die über die rein räumlichen Eigenschaften einer attendierten Konstellation hinaus, die Beziehung der Entitäten zueinander (→ Gruppierungsstärke) und/oder die kategoriale Zuordnung (→ Objektkategorisierung) voraussetzt.

Von dem präattentiven „Entdecken“ visueller Objekte durch parallele Vorverarbeitung ist die serielle attentive Verarbeitung („Erkennung“) zu unterscheiden, die den *sequentiellen* Aspekt von SFA deutlich macht. Fokussierte Aufmerksamkeit auf Objekte und deren Wechsel führen notwendigerweise zu einer Abfolge von Aufmerksamkeitszuwendungen und somit distinkten „attentional states“ (Sperling/Weichselgartner). Vielfältige Evidenz (z. B. aus dem Bereich der „Scan paths“) spricht dafür, daß diese Abfolgen im Gedächtnis abgelegt, d. h. repräsentiert, werden: „Orienting across a scene from memory need not be regarded as a two-stage-process: (i) representing, (ii) scanning. The representation itself could be directionally constructed [...]“ (Kinsbourne 1988:247). Dies entspricht der Beobachtung, daß SFA neben bottom-up- auch top-down-Aspekte (vgl. den Begriff der ‚attentional templates‘) involviert.

Obwohl Objekt- und Rauminformation auf unterschiedlichen Wegen (dem ventralen und dem dorsalen Pfad) weiterverarbeitet werden, muß die mittlerweile weit verbreitete Unterscheidung eines temporalen What- und eines parietalen Where-Systems differenziert betrachtet werden, da auch Objekte/Formen durch räumliche Eigenschaften (implizite Relationen) charakterisiert sind und nicht nur als Mengen bzgl. Rauminformation unspezifizierter Merkmale betrachtet werden können. Ein inhärentes Problem von Ansätzen, die eine solche kategoriale Objektrepräsentation-Raumrepräsentation-Distinktion durchfüh-

ren ist „that information about the spatial relationships of features within an object is lost“ (Olshausen et al. 1995:45). Die Dichotomie What/Where ist daher ungenau und in dieser Interpretation letztlich unzutreffend. Stattdessen ist Objektrepräsentation durch die Kategorisierung eines qua Aufmerksamkeitszuwendung objektbezogenen, projizierten räumlichen Ausschnitts der „map of locations“ im temporalen Bereich charakterisiert. Dabei zeigen die Untersuchungen zur Verarbeitung hierarchischer Stimuli, daß Objekterkennung *gleichzeitig* auf unterschiedlichen Granularitätsstufen stattfindet.

Raumrepräsentation ist hingegen durch die Verarbeitung der nicht-fokussierten Information sowie der Aufmerksamkeitswechsel u. a. im parietalen Bereich gekennzeichnet. Aus diesem Grund ist die Verwendung der What/Where-Dichotomie in Landau/Jackendoff (1993) zu kritisieren. Es erscheint hochgradig plausibel, daß konzeptuelle räumliche Relationen nicht durch direkten Bezug auf Raum („Where“)-Information zu charakterisieren, sondern als kategorisierte Aufmerksamkeitswechsel in unterschiedlichen räumlichen Koordinatensystemen zu betrachten sind. In jedem Fall zeigen die Erscheinungen des objektbasierten Neglekts, daß der Wechsel der Aufmerksamkeitszuwendung auf Objekte eine *notwendige* Bedingung für die Etablierung konzeptueller und für die Sprache zugänglicher räumlicher Relationen darstellt.

Ontologisch ist in Anbetracht der Forschungsergebnisse zwischen „spatiotemporal tokens“ („blobs“, „boundaries“ etc.), „object files“ und „types“ zu unterscheiden (s. auch Chun 1997). Object files weisen dabei jeweils einen oder mehrere „Indizes“ auf räumliche Token auf und „individuierten“ bzw. „instantiierten“ ihre entsprechenden konzeptuellen (Objekt-) Types. Es zeigt sich, daß die in Kapitel 2 vorgestellten Raumlogiken überwiegend räumliche Token und ihre (impliziten) Beziehungen betrachten. Zusätzlich müssen aber die Aufmerksamkeitszustände bzgl. visuell-räumlicher und konzeptueller Objekte, sowie die Wechsel solcher Zustände, berücksichtigt werden.