

# Dynamische Konzeptverarbeitung mit imaginalen und assoziativen Strukturen

Klaus Kessler, Martin Hoffhenke, Gert Rickheit und Ipke Wachsmuth

SFB-360, Universität Bielefeld, Postfach 10 01 31, D-33501 Bielefeld

## Dynamic concept processing with imaginal and associative structures

### Summary

This article presents the cooperation between an informatics and a psycholinguistic approach for modeling dynamic concept processing. As a new approach to object classification, imaginal prototypes are included in the categorical representation of the informatics framework. Objects are classified in a robust and flexible manner by gradually refined geometric shape descriptions. The psycholinguistic approach details a process model which uses the dynamics of recurrent connectionist networks to deduce and test hypotheses of temporal flow within conceptualization processes. Separation of the network into imaginal and associative subnetworks proves empirically adequate. Integration of the informatics and psycholinguistic frameworks offers a hybrid approach of dynamic concept processing.

### Zusammenfassung

Diese Arbeit stellt die Kooperation eines informatischen und eines psycholinguistischen Ansatzes bei der Modellierung dynamischer Konzeptverarbeitung vor. Im informatischen Teil werden zum Zweck formbasierter Objektklassifikation imaginale Prototypen in die kategoriale Repräsentation einbezogen. Objekte werden durch sukzessive Verfeinerung geometrischer Formmodelle flexibel und robust klassifiziert. Im psycholinguistischen Ansatz wird ein Prozeßmodell entwickelt, das die Dynamik rekurrenter konnektionistischer Netze nutzt, um Hypothesen zu zeitlichen Verläufen von Konzeptualisierungsprozessen abzuleiten und sie experimentell zu prüfen. Dabei erweist sich eine Trennung in imaginale und assoziative Subnetze als empirisch adäquat. Die Verbindung zwischen den beiden Ansätzen zielt auf einen hybriden Ansatz der Modellierung dynamischer Konzeptverarbeitung ab.

### 1 Einleitung

Die Hauptaufgabe des Teilprojekts Konzeptdynamik („CO-DY“) bei der Entwicklung eines Künstlichen Kommunikators im SFB 360 ist ein operationaler Ansatz der dynamischen Konzeptverarbeitung, welcher durch fortlaufende Aktualisierung der Wissensrepräsentation dem künstlichen Kommunikator ein möglichst optimales Verständnis der Situation in einer kooperativ bewältigten Aufgabe ermöglicht (Wachsmuth & Jung, 1996). Im Referenzszenario des SFB geht es speziell um aggregatbildende Konstruktionsaufgaben (Montagen).

Mit den bisher betrachteten strukturierten Repräsentationsformaten kann Wissen über räumliche Objektmerkmale - zum Beispiel Form und Orientierung - jedoch oft nur unvollständig modelliert werden. In der Kognitionswissenschaft besteht zunehmend Konsens darüber, daß abstrakt-strukturelle Informationen einer Ergänzung durch imaginale Information bedürfen (z.B. Engelkamp, 1983; Herrmann & Grabowski, 1994; Klix, 1980; Kosslyn, 1994; Paivio, 1979). Als Erweiterung der Konzeptmodellierung werden daher imaginale Prototypen zur Repräsentation generischer Objekte und Baugruppen integriert. Über die zusätzliche Flexibilisierung hinaus sollen an den imaginalen Prototypen weitere räumliche Merkmale, zum Beispiel intrinsische Orientierungen, verankert werden.

Diese Erweiterung entspricht einem aktuellen interdisziplinären Forschungsinteresse der Kognitionswissenschaft und Künstlichen Intelligenz an räumlichen Repräsentationen und Inferenzprozessen (z.B. Chandrasekaran & Simon, 1992; Glasgow et al., 1995). Neben der Entwicklung geometrisch basierter Formerkenner im informatischen Projektteil wurde daher in einem psycholinguistischen Teil ein Netzwerkmodell von Konzeptdynamik entworfen, um Prozeßhypothesen der menschlichen Informationsverarbeitung zu entwickeln und experimentell prüfen zu können (Kessler & Rickheit, im Druck). Die übergreifenden Fragestellungen dieser beiden Projektteile lassen sich folgendermaßen konkretisieren: *Wie ist es möglich, daß natürliche und technische kognitive Systeme Dinge an ihrer charakteristischen Form erkennen, und welche Konsequenzen ergeben sich daraus? Wird dieser Prozeß der Objektklassifikation durch den visuellen und sprachlichen Kontext beeinflusst?*

Die Untersuchung des informatischen Projektteils zielt darauf ab, Aggregate und Objekte möglichst schnell, korrekt und robust zu klassifizieren, um für den gesamten SFB ein Teilsystem bereitzustellen, welches kurzschrittig aktualisierte, möglichst umfassende dreidimensionale Repräsentationen für einen künstlichen Kommunikator bereitstellt. Der Modellierungsansatz des psycholinguistischen Projektteils ist nicht als Konkurrenz, sondern als Ergänzung zu betrachten, welche für kognitive Grundlagenforschung gemäß der experimentell-simulativen Methode (Rickheit & Strohner, 1993) konzipiert wurde und somit auch an den Kapazitätsbeschränkungen und an dem Fehlverhalten menschlicher Informationsverarbeitung interessiert ist. Aus einem Prozeßmodell werden experimentelle Hypothesen durch Simulation abgeleitet. Diese Hypothesen werden in Experimenten geprüft und entsprechende Modellanpassungen vorgenommen. Wenn sich bestimmte Modellannahmen bewähren, so dienen sie als empirische Grundlage für das informatische System, um sowohl dem technischen Anspruch als auch dem Anspruch nach kognitiver Adäquatheit gerecht zu werden. Aktuell ist die Arbeitsteilung im Projekt derart, daß im informatischen Teil ein Formerkenner für perzipierte 3D-Objekte realisiert wird, der auf imaginalen Prototypen basiert und flexible und robuste



Abb. 1. Beispiele verschiedener Ausprägungen der Kategorie Flugzeug

Objektklassifikation ermöglicht. Im psycholinguistischen Teil wird die Kontextsensitivität und die Dynamik des Prozesses der Konzeptverarbeitung modelliert und resultierende Hypothesen in Reaktionszeitexperimenten geprüft. Die beiden Ansätze werden nun nacheinander vorgestellt. Als Konsequenz dieser sich ergänzenden Arbeiten wird abschließend ein hybrides Gesamtsystem in Aussicht gestellt.

## 2 Imaginale Prototypen

Verschiedene psychologische Untersuchungen der vergangenen 20 Jahre haben gezeigt, daß der Mensch Informationen nicht nur in einer abstrakt-strukturellen Form „speichern“ und verarbeiten kann, sondern daß er häufig auch auf eine bildhafte Vorstellung von Dingen und Situationen zurückgreift. Kategorien, bzw. die Repräsentationen von Kategorien, scheinen ebenfalls nicht nur aus strukturellen oder funktionalen, sondern auch aus bildhaften Aspekten zu bestehen. Insbesondere für die Beschreibung von Kategorien durch Prototypen werden bildhafte Repräsentationen, im weiteren imaginale Prototypen genannt, postuliert (Kosslyn, 1980; Block, 1982; Kosslyn, 1994).

### 2.1 Imaginale Prototypen in der technischen Anwendung

In der ersten Projektphase des SFB 360 wurde für die Repräsentation und Erkennung von Baugruppen ein propositionaler Ansatz (COAR = Concepts for Objects, Assemblies, and Roles; s. Jung, 1997) entwickelt, der sich auf den strukturellen Aufbau eines Aggregats aus einzelnen Bauteilen stützt. Mit COAR können Baugruppen in ihrem detaillierten Aufbau beschrieben und den konstituierenden Bauteilen ihre funktionspezifische Rolle im Kontext der Baugruppe zugeordnet werden (Wachsmuth & Jung, 1996). Dieser Ansatz wird in den aktuellen Arbeiten um einen zweiten Ansatz ergänzt, der nicht die konkrete Konfiguration von Bauteilen definiert, sondern lediglich die grobe Form, die sich - unabhängig von konkreten Bauteilen - durch die Anordnung der Einzelteile ergibt. Diese Repräsentationsform nennen wir *imaginaler Prototyp* oder auch Formprototyp. Mit imaginalen Prototypen kann Wissen über die generische Form von komplexen Ob-

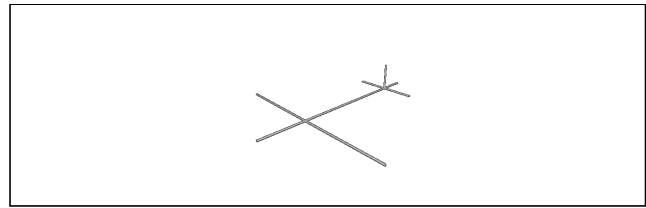


Abb. 2. Formrepräsentation eines abstrakten Flugzeugs (Skelettmodell)

jekten auf der Basis parametrischer Formmodelle beschrieben werden.

### 2.2 Repräsentation von imaginalen Prototypen

Die Verarbeitung bildhafter Repräsentationen beim Menschen ist letztendlich noch eine offene Frage. Um so mehr stellt sich die Frage, wie ein technisches System bildhafte Repräsentationen mit ihren unterschiedlichen Anforderungen darstellen kann. Dazu sollte zunächst eingegrenzt werden, was die (technischen) imaginalen Prototypen zu leisten haben. Dies sind im wesentlichen zwei konfligierende Anforderungen. Es muß einerseits möglich sein, Form in ihrer abstraktesten Weise zu repräsentieren. Andererseits muß es auch möglich sein, durch die Repräsentation von diskriminierenden Formmerkmalen Unter- von Oberkategorien zu differenzieren. Imaginale Prototypen müssen also unterschiedlich abstrakte Kategorien beschreiben können.

In der Literatur findet man verschiedene Ansätze zur (technischen) Repräsentation von Form. Einige Arbeiten beschränken sich ausschließlich auf eine 2D-Repräsentation (Blum & Nagel, 1978; Siddiqi et al., 1998), die Tendenz geht jedoch überwiegend zu 3D-Repräsentationen (Marr & Nishihara, 1978; Biederman, 1987), auch im Bereich der Bilderkennung (Brooks, 1984; Ferryman et al., 1995), da 3D-Modelle eine betrachterstandpunkt-invariante Repräsentation erlauben<sup>1</sup>. Diese Arbeiten beschränken sich jedoch immer auf eine singuläre Repräsentationsform und erlauben keine Modellierung kategorienadäquat abstrahierter Formprototypen. Sie bieten damit kein universelles Modell eines imaginalen Prototypen, das die oben genannten gegensätzlichen Forderungen erfüllen könnte. Wir postulieren daher einen integrierten Ansatz gradiert abstrahierter Prototyparten.

Die *erste Anforderung* an einen imaginalen Prototypen lautete, Form in ihrer abstraktesten Weise zu repräsentieren. Dies läßt sich gut anhand des Beispiels der Kategorie Flugzeug veranschaulichen, deren Entitäten durch vielfältige Formvarianten gekennzeichnet sind (s. Abb. 1). Es stellt sich

1. Es sei darauf hingewiesen, daß die Frage nach der betrachterstandpunkt-invarianten Repräsentation beim Menschen noch nicht geklärt ist (siehe z.B. Tarr & Bühlhoff (1998) für den betrachterabhängigen bzw. Hummel (1998) für den betrachterunabhängigen Standpunkt). In der technischen Anwendung bietet die invariante Repräsentation Vorteile. Insbesondere in dem hier vorgestellten Projekt liegt als Perzept eine dreidimensional rekonstruierte Szenenbeschreibung vor; eine dreidimensionale Modellierung der Prototypen liegt daher nahe.

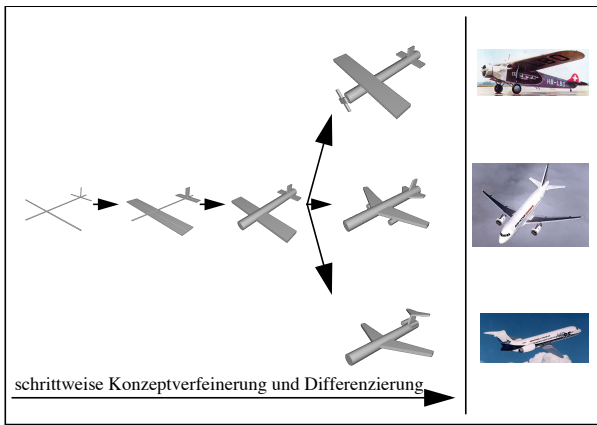


Abb. 3. Kaskadierte Anordnung imaginaler Prototypen: Skelettmodell, Flächenstrukturmodell, Volumenmodell

die Frage, welche charakteristische Form die Kategorie Flugzeug beschreibt. Insbesondere das Flugzeugmodell aus den Baufixteilen (s. Abb. 1 unten) macht sehr schnell deutlich, daß es sich nicht um detaillierte Formmerkmale handelt, die für ein Flugzeug typisch sind. Trotz fehlendem Tragflächenprofil, Löchern in Tragflächen und Rotorblättern und untypischen Erhebungen auf der Oberseite ist das Modell für einen Menschen sofort und eindeutig als „Flugzeug“ zu erkennen. Es sind die räumlichen Ausdehnungen, die sich in einer konkreten Konfiguration zueinander befinden, die das Objekt als Flugzeug charakterisieren. Das an eine Strichzeichnung erinnernde dreidimensionale *Skelettmodell* (auch *Längenstrukturmodell*), als abstrakteste aller möglichen Darstellungsformen, bietet hierfür ein adäquates Beschreibungsmittel (s. Abb. 2).

Die *zweite Anforderung* an einen imaginalen Prototypen ist die Möglichkeit der Modellierung diskriminierender Formmerkmale. Dies ist durch ein Skelettmodell allein nicht zu leisten, da eine Differenzierung meist erst in der zwei- oder dreidimensionalen Ausprägung der Formmerkmale hervortritt (s. beispielsweise die charakteristische Deltaform der Jagdflugzeugtragflächen in Abb. 1). Benötigt werden demnach auch Methoden zur Beschreibung höherdimensionaler Formmerkmale. Damit erscheint also eine Integration von Beschreibungsmechanismen unterschiedlicher Expressivität zu einem Gesamtkonzept eines imaginalen Prototypen notwendig.

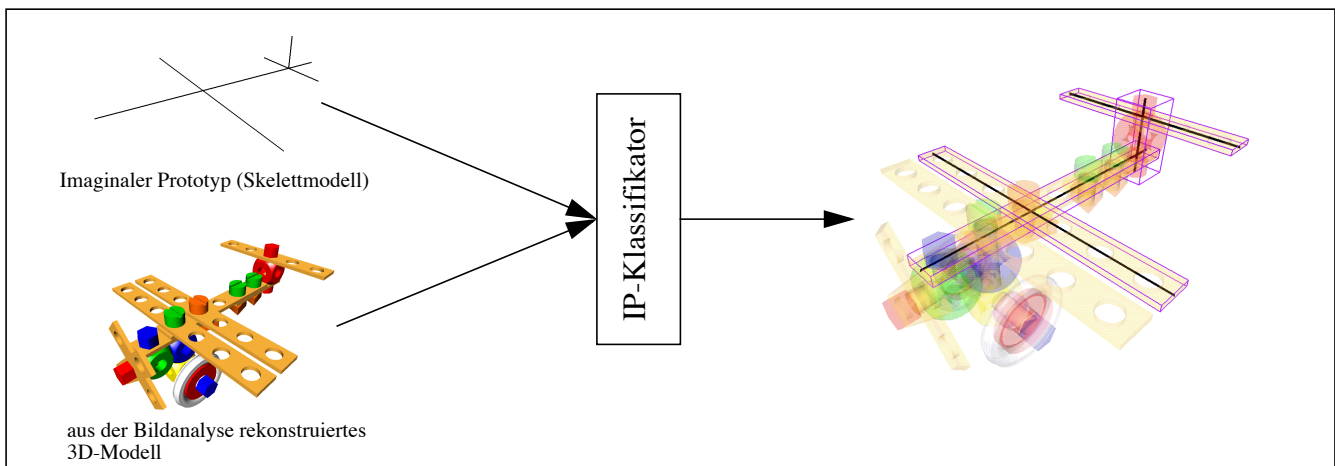


Abb. 4. Klassifikation des Baufixflugzeugs gegen den imaginalen Prototypen (IP) von Flugzeug

### 2.3 Integration der Prototypen in einer Kaskadenarchitektur

Bei der Integration der einzelnen Prototyparten sind zwei wesentliche Punkte zu berücksichtigen. Erstens, das Einbetten der Prototypen in eine einheitliche Konzeption, um eine Interoperabilität zwischen diesen zu erreichen, und zweitens die Bewahrung der jeweiligen Eigenarten der einzelnen Prototyparten, um ihre spezifische Expressivität zu erhalten. Dabei ist zu beachten, daß sich die abstrakte Ausdruckskraft eines Prototypen nicht nur in seinem (geometrischen) Formmodell, sondern auch in dem auf diesem operierenden Klassifikationsalgorithmus manifestiert.

Um die individuellen Charakteristiken jeder Prototypart zu wahren, werden sie einzeln, jede für sich mit ihrem zugehörigen Klassifikationsalgorithmus, in eine Umgebung mit uniformen Ein-/Ausgabestrukturen eingebettet und kaskadenartig mit abnehmendem Abstraktionsgrad angeordnet (s. Abb. 3), d.h. es wird eine zunehmende Formdifferenzierung berücksichtigt. Diese findet Niederschlag in der Kaskadenarchitektur des Formerkenners.

Der Klassifikationsprozeß beginnt, indem das Perzept<sup>2</sup> zunächst mit dem Prototypen der höchsten Abstraktionsstufe abgeglichen wird. Konnte hierbei eine erfolgreiche Zuordnung von Formmerkmalen stattfinden (wie in Abb. 4 veranschaulicht), wird eine Hypothese generiert, die dem Abgleich des Objekts mit dem Prototypen des nächst geringeren Abstraktionsgrads als Ausgangsinformation dient. Kann die Hypothese auch auf dieser Ebene bestätigt werden, wird eine neue Hypothese mit angereicherter Information generiert und der folgenden Ebene übergeben.

Die Abb. 5 zeigt das Prinzip des Klassifikationsablaufs am Beispiel des Flugzeugs auf den drei kaskadierten Ebenen der Längenstrukturmodelle (L), Flächen- (F) bzw. Volumenmodelle (V). Nachdem durch einen Montageschritt das Aggregat Aggregat-0001 entstanden ist, wird daraus zunächst eine allgemeine Null-Hypothese generiert, d.h. die Hypothese ist noch durch keinen Klassifikationsvorgang bestätigt worden. Für die Hypothese existieren auch noch keine Annahmen über den möglichen IP-Konzepttyp. Daher werden zunächst alle IP-Modelle getestet. Das Flugzeugmodell wird im ge-

2. Im Montage-Szenario des SFB liegt dem künstlichen Kommunikator als Perzept eine dreidimensionale geometrische Rekonstruktion der aktuellen Szene vor.

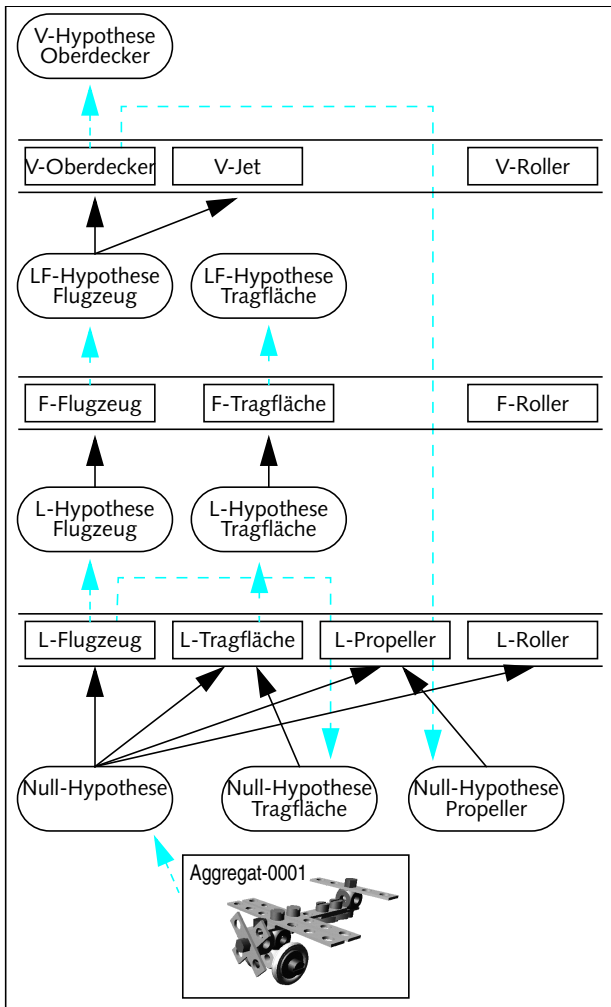


Abb. 5. Klassifikation mit den kaskadenförmig angeordneten imaginalen Prototypen. Die Rechtecke symbolisieren die Formmodelle, die Ovale die generierten Hypothesen. Die durchgezogenen Pfeile zeigen einen Abgleich zwischen Hypothese und Formmodell an, die gestrichelten Pfeile die Erzeugung einer neuen Hypothese.

zeigten Beispiel als einziges als passend erkannt. Daher wird daraus eine Hypothese der nächst höheren Ebene (L-Ebene) generiert. Als Bestandteil wurde unter anderem die Tragfläche identifiziert. Da sie einem eigenen IP-Konzept zugeordnet ist, wird aus ihr eine eigene Hypothese erstellt, zunächst wieder auf der Null-Ebene. Hier ist schon die Annahme einer Tragfläche bekannt. Daher ist hier dann auch nur das IP-Modell der Tragfläche passend. Zur selben Zeit, zu der die Tragflächen-Hypothese auf der L-Ebene getestet wird, wird auf der F-Ebene nebenläufig die Flugzeug-Hypothese gegen das einzig passende IP-Modell, das Flugzeugmodell, getestet. Dabei können die auf der L-Ebene gewonnenen Informationen über die Rolle bestimmter Bestandteilekomponenten aufgegriffen werden und auf diese Weise die Klassifikation deutlich vereinfachen. Die Hypothese kann bestätigt werden und es wird wiederum eine neue Hypothese (diesmal auf der F-Ebene) generiert. Diese paßt auf die IP-Modelle für den Oberdecker und den Jet, da beide Spezialisierungen vom Flugzeug sind. Der Oberdecker stellt sich im Beispielfall als das zutreffende Modell heraus. Da Oberdecker einen Propeller besitzen, wird auch hier wieder eine entsprechende Bestandteil-Hypothese auf der Null-Ebene generiert.

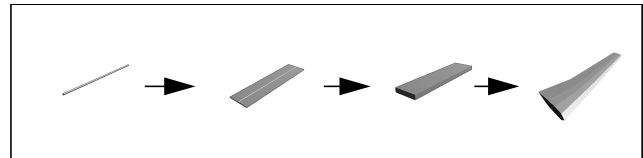


Abb. 6. Die Ausdrucksfähigkeit eines Prototypen für ein gering strukturiertes Objekt wie die Tragfläche liegt bei einem deutlich geringeren Abstraktionsniveau

Jede Hypothese ist mit einer Gütebewertung (genannt Salienz) versehen. Dieser Salienzwert setzt sich aus verschiedenen Faktoren zusammen, wie beispielsweise der *Aktivierung* der entsprechenden Kategorie (s. auch Abschnitt 3) und der *Signifikanz* des jeweiligen Prototypen. Die Signifikanz stellt einen Wert dar, mit dem die Ausdruckskraft des jeweiligen Prototypen bewertet wird. Dies veranschaulicht die Abb. 6. Der Abstraktionsgrad eines aussagekräftigen Prototypen liegt bei einem gering strukturierten Objekt wie der Tragfläche deutlich niedriger als bei einem stark strukturierten Objekt, wie es das Flugzeug in Abb. 3 darstellt.

Weitere Beispiele von als Flugzeug klassifizierten Baugruppen zeigt die Abb. 7. Der bereits vollständig implementierte Skelettmodell-Klassifikator benötigt für die Erkennung der gezeigten und ähnlicher Flugzeug-Modelle bis zu 5 Sekunden auf einer Sun ULTRA 10. Vom kognitiven Standpunkt unbefriedigend ist dabei, daß die Klassifikationszeit noch von der Komplexität der Perzepte abhängt. Hier wäre eine deutliche Verbesserung durch eine perzeptiv gruppierende Vorverarbeitung zu erwarten. Ebenfalls vom kognitiven Standpunkt ist allerdings bereits jetzt positiv zu vermerken, daß mit dem Formerkennung die Generalisierung auf verschiedene Ausprägungen eines Objekttyps erfolgreich bewältigt wird und zudem (durch stark abstrahierte Formbeschreibung) eine Laufzeitverbesserung gegenüber der Klassifikation mit dem strukturellen Beschreibungsformalismus COAR erzielt wird.

#### 2.4 Anreicherung von imaginalen Prototypen mit räumlichen Attributen

Ein weiterer Aspekt, der die Verwendung von imaginalen Prototypen gegenüber propositionalen Beschreibungen motiviert, ist die einfache Anreicherung des Prototypen mit räumlichen Relationen wie beispielsweise einer intrinsischen Orientierung des Objekts. Diese Anreicherung kann einfach über das Hinzufügen geometrischer Objekte erfolgen. So können beispielsweise Quader oder Kugeln Regionen mit einer bestimmten Bedeutung kennzeichnen oder Vektoren ausgezeichnete Richtungen wie beispielsweise ‚vorne‘

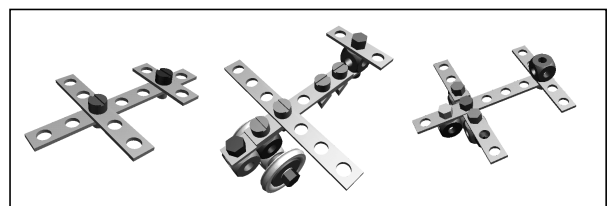


Abb. 7. Weitere Beispiele von als Flugzeug erkannten Baugruppen.



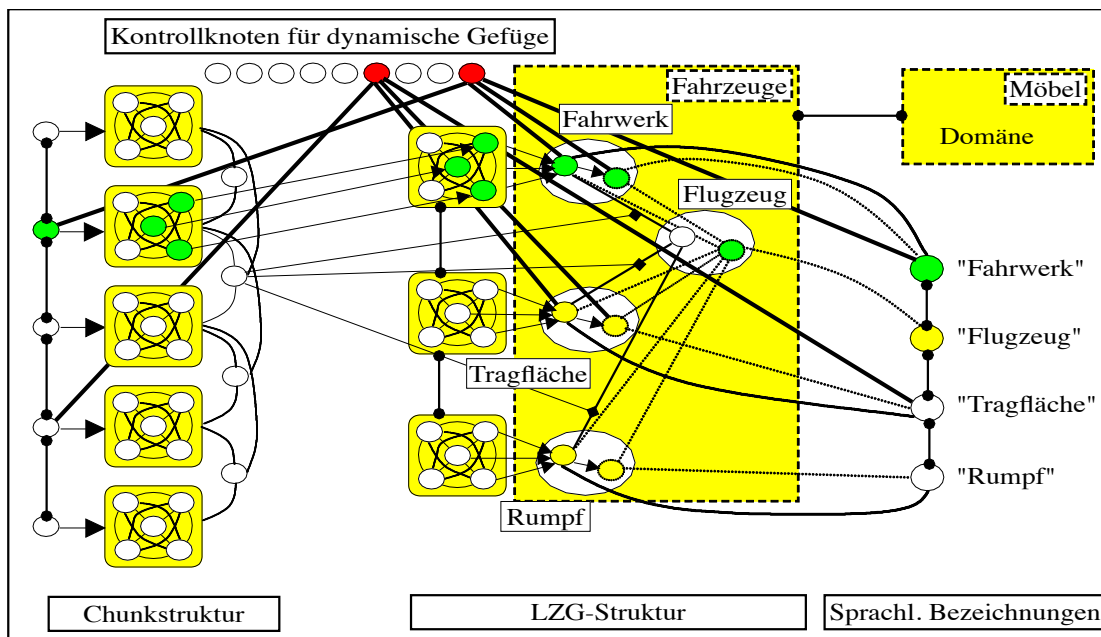


Abb. 8. Das Modell von Kessler und Rickheit (im Druck). Pfeile bezeichnen die Richtung der Aktivationsausbreitung, Linien stehen für Verbindungen in beide Richtungen, Linien mit Punkten am Ende sind inhibitorisch, Linien mit einem Rhombus am Ende sind multiplikativ. Die dynamischen Verbindungsgefüge sind durch dicke durchgehende Linien dargestellt, welche auf jeweils einen Kontrollknoten konvergieren. Innerhalb der LZG-Struktur bedeuten gepunktete Linien assoziative und durchgehende Linien imaginale Substrukturen des Langzeitgedächtnisses (LZG).

definieren. Derartige Informationen bilden die Basis, um sprachliche Konstruktionen wie: „stecke den Propeller vorne an das Flugzeug“ zu einer konkreten Handlung auflösen zu können.

### 2.5 Weitere Arbeiten

Ein wesentlicher Untersuchungsgegenstand weiterer Arbeiten wird es sein, wie sich die Aspekte von propositionalen und imaginalen Anteilen sinnvoll ergänzen können. Solche Interaktionen werden beispielsweise benötigt bei Anweisungen wie der zuvor erwähnten: „stecke den Propeller vorne an das Flugzeug“. Die imaginalen Prototypen können eine ungefähre Ziel-Lage des Propellers beschreiben, für den konkreten Montageschritt wird dann aber zusätzlich das Wissen über die Fügemöglichkeiten der verwendeten Bauteile benötigt, wie es mit COAR (s. Abschnitt 2.1) zur Verfügung steht. Ein Thema wird auch sein, wie die beiden Repräsentationsformate bei der Rollenzuschreibung an einzelne Bauteile zusammenwirken können.

### 3 Eine Netzwerkkonzeption von „Konzeptdynamik“

Die Netzwerkkonzeption basiert auf den vorangegangenen Überlegungen zur Objekterkennung und greift ebenfalls die Grundidee eines Form-basierten Ansatzes für die Modellierung einer imaginalen Szenenrekonstruktion auf. Somit existiert eine Vorstellung darüber, wie Baugruppen aus einer visuell zugänglichen Konstruktionsszene in einem Netzwerkmodell abstrahiert werden können. Das hier vorzustellende Modell ist im Gegensatz zu den eher technischen Überlegungen des vorangegangenen Abschnitts dazu konzipiert wor-

den, wichtige Prozeßeigenschaften menschlicher Informationsverarbeitung nachzuvollziehen, um experimentelle Hypothesen insbesondere zu Reaktionszeitverläufen ableiten zu können. Eine wichtige Annahme betrifft die Anzahl der Informationseinheiten, welche gleichzeitig und distinkt in einer Netzwerkstruktur aufrechterhalten werden können. Mehrere perzipierte Objekte als Input können im Netz die verschiedensten Units aktivieren, ohne daß man anschließend feststellen kann welche Unit durch welches Objekt aktiviert wurde (vgl. z.B. Treisman & Schmidt, 1982; Fodor & McLaughlin, 1990). Es bedarf folglich zusätzlicher Mechanismen, die Variablen (aktivierte Units) eines Objekts als repräsentierte Einheit zusammenzuhalten (vgl. Fodor & McLaughlin, 1990; Goebel, 1991). Diese zusätzlichen Mechanismen können durch Kontrollknoten oder oszillierende Units implementiert werden (z.B. Smolensky, 1987; Shastri & Ajjanagadde, 1990). Diese Mechanismen zeichnen sich jedoch dadurch aus, daß der benötigte Aufwand (entweder an Kontrollunits oder an Oszillationsphasen) mit der Zahl der gleichzeitig zu repräsentierenden Einheiten ansteigt (vgl. Fodor & McLaughlin, 1990). Solche Mechanismen lassen folglich nur in bezug auf wenige mentale Einheiten effektiv anwenden. Dies deckt sich jedoch gut mit Ergebnissen zur Kapazitätsbeschränkung des Arbeitsgedächtnisses, wonach dem Menschen 7 solcher Einheiten (Chunks) zur Verfügung stehen (Miller, 1956). Dabei ist die Anzahl der Chunks nicht der wesentliche Punkt sondern die Tatsache der Begrenzung selbst. Aufbauend auf diesen Befunden werden im Modell fünf Chunks angenommen. Die Chunks besitzen eine Formstruktur, die auf der Modellierung im informatischen Teil basiert und welche der Formstruktur der Konzepte im Langzeitgedächtnis (LZG) entspricht, so daß die Chunks sequentiell mit dem LZG abgeglichen werden können. Dabei wird ein Kategorisierungsprozeß zum einen durch bereits ka-

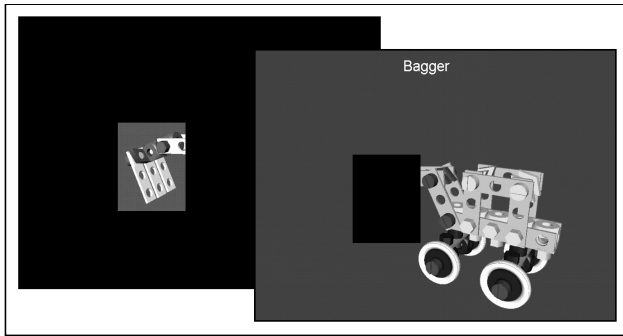


Abb. 9. Ein Beispiel für ein erstes (links) und ein zweites Bildschirmbild (rechts) in Experiment 1. Die Baggerschaufel (links) ist erkennbar mit einem verdeckten, größeren Objekt strukturell verbunden.

tegorisierten visuellen Kontext und zum andern durch sprachliche Information beeinflusst. Nachdem das Netz einen stabilen Zustand erreicht hat, werden Chunks und aktivierte LZG-Konzepte über Kontrollknotenstrukturen dynamisch verknüpft<sup>3</sup> (Abb. 8).

Zur Modellierung von Konzept-Zusammenhängen sind die LZG-Knoten untereinander vernetzt. Zum aktuellen Modellierungsstand werden zwei Arten von Konzeptzusammenhängen berücksichtigt, zum einen Teil-Ganzes-Relationen zwischen Fahrzeugteilen und dem entsprechenden Gesamtfahrzeug, welche durch exzitatorische Verbindungen zwischen Fahrzeug- und Aggregatknotten repräsentiert werden, und zum anderen Domänen-Relationen, welche Konzepte in verschiedenen Domänen inhibitorisch verknüpfen. Des weiteren gibt es zwei verschiedene, partiell unabhängige Substrukturen im LZG (Abb. 8; vgl. Herrmann et al., 1996). Es bestehen erstens imaginaler Verbindungen über ein Subnetz, welches Fahrzeuge und ihre Teile verbindet. Diese Konnektionen leiten nur dann Aktivierung weiter, wenn Fahrzeugteile tatsächlich räumlich-strukturell mit einem Gesamtobjekt verbunden sind (z.B. wenn eine Baggerschaufel tatsächlich auf einen Bagger montiert ist, so aktiviert das LZG-Konzept #Schaufel# das LZG-Konzept #Bagger# über die imaginalen Verbindungen). Zweitens sind assoziative Verbindungen vorhanden, welche immer Aktivierung weiterleiten, gleich in welcher räumlich-strukturellen Beziehung Teil und Ganzes zueinander stehen (z.B. wenn eine Baggerschaufel nicht an einen Bagger montiert ist, so aktiviert #Schaufel# das Konzept #Bagger# über die assoziativen, nicht jedoch über die imaginalen Verbindungen). Die Annahme dieser beiden Strukturen hat wesentliche Auswirkungen auf die Aktivationsdynamik im Modell. Diese Modellannahmen wurden zu Hypothesen konkretisiert und in drei Experimenten<sup>4</sup> der empirischen Prüfung ausgesetzt (für eine detaillierte Darstellung der Experimente und Ergebnisse verweisen wir auf Kessler & Rickheit, im Druck).

3. Für eine detailliertere Beschreibung dieses Modellaspekts verweisen wir auf (Kessler & Rickheit, im Druck).

4. Das Stimulusmaterial in den Experimenten wurde mit dem CODY Virtuellen Konstrukteur des informatischen Projektteils erstellt (Jung et al. 1998). Es umfaßte ein Flugzeug, einen Bagger, ein Motorrad und eine Lokomotive.

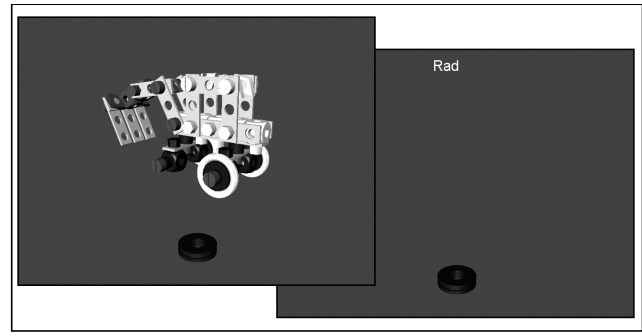


Abb. 10. Ein Beispiel für ein erstes (links) und ein zweites Bildschirmbild (rechts) in Experiment 2.

### 1. Experiment

In diesem Experiment wurde die Hypothese geprüft, daß die Aktivationsausbreitung im LZG in zwei Subnetzen stattfindet: in einem imaginalen und in einem assoziativen Subnetz. Nur bei struktureller Verbundenheit zwischen zwei Chunks (s.o.) werden die imaginalen Verbindungen freigeschaltet, sind dann aber effektiver als die permanent verfügbaren assoziativen Verbindungen. In einem ersten Bild wurde ein maskierter Stimulus mit einem sichtbaren Ausschnitt auf dem Bildschirm gezeigt (s. Abb. 9). Der Ausschnitt stellt ein Fahrzeugteil dar (z.B. Baggerschaufel), das entweder erkennbar räumlich-strukturell mit „etwas“ verbunden war oder sichtbar räumlich-strukturell isoliert war. Bei einer räumlich-strukturellen Verbindung ist das Gesamtobjekt (#Bagger#) laut Modell voraktiviert und somit sollte die Kategorisierung und folglich auch die Benennung des Objekts schneller verfügbar sein als bei Darbietung eines isolierten Fahrzeugteils. Auf einem zweiten Bild wurde der maskierte Teil des ersten Bildes gezeigt und konnte ein passendes bzw. ein unpassendes Fahrzeug darstellen (vgl. Abb. 9). Bei der Darbietung des zweiten Bildes wurde per Reaktionszeitmessung die Verfügbarkeit der Kennzeichnung gemessen.

### Die wichtigsten Ergebnisse

Die Annahme zweier partiell unabhängiger Subnetze in der LZG-Struktur (s.o.) konnte durch die hier vorgestellten Befunde bestätigt werden. Allein durch die Andeutung einer räumlich-strukturellen Verbindung eines Fahrzeugteils in einem ansonsten maskierten Stimulus konnte das gesamte Fahrzeug effektiver voraktiviert werden als bei einer isolierten Darbietung des ansonsten identischen Fahrzeugteils. Dies ist eine Bestätigung der Annahme, daß das imaginale Netz nur unter bestimmten perzeptuellen Bedingungen genutzt wird, dann jedoch eine höhere Effektivität als das assoziative Teilnetz aufweist. Evidenz für das Vorhandensein des assoziativen Netzes kann aus dem Befund abgeleitet werden, daß ein isoliert dargebotenes passendes Fahrzeug schneller mit einer Kennzeichnung abgeglichen werden kann als ein isoliert dargebotenes, jedoch falsches Fahrzeug. Bei letzterem bestehen keine assoziativen Verbindungen zum dargebotenen maskierten Fahrzeugteil.

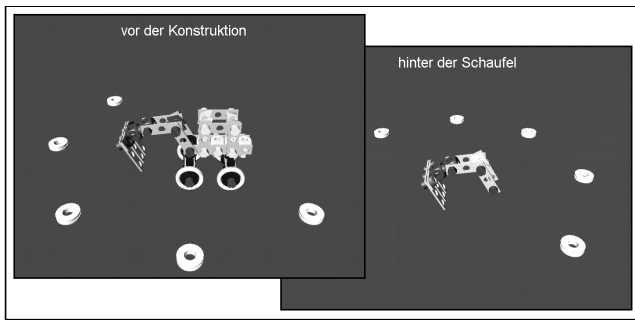


Abb. 11. Zwei Beispielstimuli aus Experiment 3. Links ist das Relatum ein ganzes Fahrzeug, die Bezeichnung des Relatums ist unspezifisch („Konstruktion“), und die Lokalisation erfolgt anhand der Präposition „vor“. Rechts ist das Relatum ein gerichtetes Fahrzeugteil, die Bezeichnung des Relatums ist spezifisch („Schaufel“), und die Lokalisation erfolgt anhand der Präposition „hinter“.

## 2. Experiment

Die Kategorisierung eines Chunks wird durch den visuellen und durch den sprachlichen Kontext beeinflusst. Ein Objekt - beispielsweise eine Leiste - kann sowohl in der Flugzeugdomäne als „Tragfläche“ als auch in der Bauteildomäne als „Leiste“ kategorisiert werden. Welche Bedeutung dominiert, hängt dabei unter anderem vom visuellen Kontext ab. Wenn mehrere Bauteile eines Flugzeugs bereits erkannt wurden (z.B. Rumpf, Fahrwerk und Propeller), so wird über die assoziativen Verbindungen das Konzept Flugzeug aktiviert, welches wiederum Aktivierung an die Konzeptknoten noch fehlender Flugzeugteile wie Tragfläche weiterleitet. Dadurch ist in diesem Kontext die Tragflächen-Bedeutung einer Leiste mindestens gleich stark wie „Leiste“. In einem bedeutungslosen Objektkontext wird hingegen eine Dominanz der Bauteilbedeutung „Leiste“ erwartet. Dieser Effekt sollte sich mindestens ebenso stark finden lassen, wenn der Kontext nicht durch Fahrzeugteile, sondern durch das Fahrzeug selbst konstituiert wird.

Zuerst wurde ein Bild dargeboten, welches ein Bauteil in einem bestimmten Kontext zeigt, dann wurde das Bauteil in einem zweiten Bild ohne Kontext, jedoch zusammen mit einer Kennzeichnung eingeblendet (s. Abb. 10). Das Bauteil im Vordergrund konnte in einem bedeutungsvollen Kontext eine Fahrzeugfunktion übernehmen. In Abb. 10 kann die Scheibe beispielsweise die Funktion des fehlenden Baggerrades übernehmen. Damit wurde es ermöglicht, daß die Scheibe im Kontext von #Bagger# mit dem Konzeptknoten von #Rad# verknüpft werden konnte. Ob dies der Fall war, wurde durch eine anschließende Kennzeichnungsbeurteilung geprüft.

### Die wichtigsten Ergebnisse

Es konnte gezeigt werden daß der visuelle Objektkontext massiven Einfluß auf die Kategorisierung eines Bauteils hatte. Bei einem sinnlosen Kontext dominiert die Bauteilkategorisierung eine Fahrzeugkategorisierung hinsichtlich ihrer Verfügbarkeit, was sich in der schnelleren Reaktion auf die

Bauteilbezeichnung ausdrückt. Dies ändert sich bei einem sinnvollen Kontext, wo sich die beiden Kategorisierungsvarianten nicht mehr statistisch bedeutsam unterscheiden. Die Fahrzeugkategorisierung weist dabei eine signifikante Verbesserung der Verfügbarkeit gegenüber einem sinnlosen Kontext auf. Dieses Ergebnis zeigt sich sowohl beim Fahrzeug- als auch beim Fahrzeugteil-Kontext.

## 3. Experiment

In diesem Experiment wurden zwei Hypothesen überprüft. Erstens, die Kategorisierung eines Chunks wird durch den sprachlichen Kontext beeinflusst. Bei gleichem visuellem Input (Fahrwerk) kann unterschiedlich spezifische sprachliche Information („Fahrwerk“ vs. „Baugruppe“) zu unterschiedlichen aktuellen Repräsentationen führen. Einmal dominiert die Flugzeugbedeutung, einmal dominiert die Bauteilbedeutung. Zweitens stellt die dynamische Verbindung eines Chunks mit einem LZG-Konzept kategoriale Information beispielsweise über die intrinsische Gerichtetheit des Objekts bereit, welche für die Rezeption von Lokalisationsäußerungen im Konstruktionsdialog genutzt werden kann. Im Modell wird von einem Ökonomieprinzip der LZG-Organisation ausgegangen (vgl. z.B. Smith, Shoben & Rips, 1974), so daß nur die LZG-Konzepte von gesamten Fahrzeugen Gerichtetheit direkt aktivieren können. Die LZG-Konzepte von Teilen können diese Information nur über das assoziative Netz aktivieren, es dauert folglich länger. Bei nicht eindeutig gerichteten Teilen, welche nur eine Achse ohne eindeutige Vorder- und Rückseite aufweisen (z.B. Fahrwerk), dauert der Prozeß durch die Uneindeutigkeit der Information noch länger. Bei ungerichteten Teilen ist diese Information überhaupt nicht verfügbar.

Es wurde ein Lokalisationsexperiment durchgeführt, wobei die Struktur und die sprachliche Bezeichnung der Relata variiert wurden (vgl. Abb. 11). Das Relatum ist das Objekt, in bezug auf welches sprachlich lokalisiert wird. Die Struktur der Relata hatte vier Ausprägungen: ganzes Fahrzeug, gerichtetes Fahrzeugteil, geachstes Fahrzeugteil (s.o.) und ungerichtetes Fahrzeugteil. Die Bezeichnung des Relatums konnte spezifisch oder unspezifisch sein („Bagger“ vs. „Konstruktion“). Um das Relatum herum war ein Halbkreis mit Scheiben angeordnet (s. Abb. 11). Die Lokalisationsaufgabe für die Vpn bestand darin, aufgrund einer sprachlichen Anweisung wie „Nehmen Sie die Scheibe vor der Konstruktion“ eine der Scheiben auszuwählen, indem sie die Scheibe mit der Maus anklickte. Es wurden hierbei die beiden Präpositionen „vor“ und „hinter“ verwendet. Der Halbkreis aus Scheiben war in jedem Fall so angeordnet, daß ein Konflikt zwischen einer egozentrischen Perspektive („Von mir aus gesehen vor dem Bagger“) und der intrinsischen Perspektive („Vor der Vorderseite des Baggers“) existierte.

### Die wichtigsten Ergebnisse

Die Struktur des Relatums hatte die erwartete Auswirkung auf die Wahl einer Perspektive, was sich in einer kontinuierlichen Verlagerung von primär intrinsischen Lokalisationen bei einem gerichtetem Fahrzeug hin zu einer Dominanz der

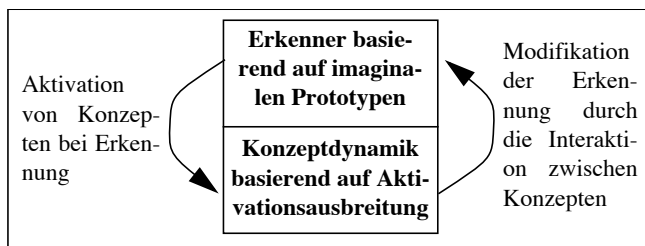


Abb. 12. Interaktionsschema des psycholinguistischen und des informatischen Projektteils

egozentrischen Perspektive bei einem ungerichteten Relatum ausdrückt. Essentiell für die Modellprüfung ist, daß die Bedeutung des sprachlichen Kontextes für die Generierung einer aktuellen Repräsentation nachgewiesen werden konnte. Eine spezifische Kennzeichnung bewirkt - bei ansonsten völlig identischem Stimulus - eine Bevorzugung der intrinsischen Perspektive. Die Auswertung der Perspektivenwahl und der Reaktionszeiten bestätigt die Annahme eines Ökonomieprinzips bei der Organisation kategorialen Wissens. Bei einem ganzen Fahrzeug war die Gerichtetheit am häufigsten (Perspektivenwahl) und am schnellsten (Reaktionszeiten) verfügbar. Bei gerichteten und geachteten Teilen wurde dennoch häufiger intrinsisch interpretiert als bei ungerichteten Relata, was insgesamt für ein Ökonomieprinzip spricht, da Gerichtetheit durch Teile aktiviert werden kann, jedoch nicht im gleichen Ausmaß wie durch das gesamte Fahrzeug.

#### 4 Zusammenfassung und Ausblick

In dieser Arbeit wurde ein umfassender Ansatz zur Aufklärung und Realisierung dynamischer Konzeptverarbeitung in natürlichen und technischen kognitiven Systemen beschrieben. Es wurde einerseits ein Entwurf zur operationalen Repräsentation imaginaler Prototypen in einem technischen System vorgestellt, welches eine flexible und robuste Objektklassifikation durch geometrische Formerkennung ermöglicht. Dabei leistet der Erkenner eine Generalisierung auf verschiedene Ausprägungen eines Objekttyps. Andererseits wurde ein Prozeßmodell dynamischer Konzeptverarbeitung entwickelt, welches die Ableitung experimenteller Hypothesen zu Reaktionszeitverläufen ermöglichte. Es konnte experimentell gezeigt werden, daß die Generierung einer aktuellen Repräsentation unmittelbar von imaginalen und assoziativen Strukturen abhängt, daß die dynamische Aktivierung von Konzeptstrukturen sensitiv gegenüber visuellem und sprachlichem Kontext verläuft und daß die Zuschreibung von intrinsischen Objekteigenschaften Teil einer dynamischen Konzeptverarbeitung mit imaginalen und assoziativen Strukturen ist.

Ein technisches Ziel des Projekts ist die Erstellung eines laufenden Prototypen des imaginalen Formerkenners, welcher die Ergebnisse der psycholinguistischen Experimente berücksichtigen und integrieren soll. Hierfür ist eine Kopplung der beiden entwickelten Systeme in Planung, wie sie in Abb. 12 angedeutet wird. Der auf imaginalen Prototypen basierende Objekterkenner wird bei der Bewertung seiner Hypothesen die Aktivierung der entsprechenden Kategorien, wie sie das Prozeßmodell des psycholinguistischen Teilprojekts liefert, berücksichtigen und bei erfolgreicher Erkennung von

Objekten entsprechende Gedächtniskonzept-Knoten aktivieren. Auf diese Weise soll eine dynamische Konzeptverarbeitung von Bauteilen und Aggregaten in einem technischen System verwirklicht werden. Als weiteres Ziel des Projekts wird das Prozeßmodell weiterentwickelt, um konform dem Anspruch nach Grundlagenforschung weitere Prinzipien menschlicher Informationsverarbeitung experimentell-simulativ aufzudecken, die u.a. die Integration multimodaler Konzeptanteile und die Dynamik ihrer Interaktion betreffen, und wiederum die Möglichkeit bieten, Anhaltspunkte für verfeinerte kommunikative Fähigkeiten technischer Systeme zu erlangen.

In der zukünftigen Arbeit soll die Planung von Handlungen mit einbezogen werden, da sie sich ganz natürlich aus den Eigenschaften der gewählten multiplen Repräsentationsform ergibt: Aus den strukturellen (COAR-)Repräsentationen ergeben sich Rollenzuschreibungen an einzelne Bauteile, aus den imaginalen Prototypen lassen sich intrinsische Objektorientierungen ableiten, und über die imaginal-assoziativen Konzeptverknüpfungen können übergeordnete Konzeptknoten aktiviert werden. Alle diese Repräsentationseigenschaften können zur weiteren Planung von Objektmanipulationen genutzt werden.

#### Literatur

- Biederman, I. (1987). Recognition-by-components: a theory of human image understanding. In *Psychological Review*, 94 (2), 115 - 147.
- Block, N. (Ed.) (1982). *Imagery*. MIT Press, Cambridge (MA).
- Blum, H., Nagel, R. N. (1978). Shape description using weighted symmetric axis features. In *Pattern Recognition*, 10, 167 - 180, Pergamon Press Ltd.
- Brooks, R. A. (1984). *Model-based computer vision*. UMI Research Press, Ann Arbor, Michigan.
- Chandrasekaran, B. & Simon, H. (1992). *Reasoning with diagrammatic representations*. Technical Report SS-92-02. Menlo Park (CA): AAAI Press.
- Engelkamp, J. (1983). Word meaning and word recognition. In T. B. Seiler & W. Wannemacher (Eds.), *Concept development and the development of word meaning*. Berlin: Springer.
- Ferryman, J. M., Worrall, A. D., Sullivan, G. D. & Baker, K. D. (1995). A generic deformable model for vehicle recognition. *Proceedings of British Machine Vision Conference*, University of Birmingham, 127-136.
- Fodor, J & McLaughlin, B. P. (1990). Connectionism and the problem of systematicity: Why Smolensky's solution doesn't work. *Cognition*, 35, 183-204.
- Glasgow, J., Narayanan, N. & Chandrasekaran, B. (1995). *Diagrammatic reasoning: cognitive and computational perspectives*. Menlo Park (CA): AAAI Press.
- Goebel, R. (1991). Binding, episodic short-term memory, and selective attention, or why are PDP models poor at symbol manipulation. In D.S. Touretzky, J.L. Elman, T.J. Sejnowski & G.E. Hinton (Hrsg.). *Connectionist models. Proceedings of the 1991 Summer School*, 253-264, San Mateo: Morgan Kaufmann.



- Herrmann, Th. & Grabowski, J. (1994). *Sprechen, Psychologie der Sprachproduktion*. Heidelberg: Spektrum Akademischer Verlag.
- Herrmann, Th., Grabowski, J., Schweizer, K. & Graf, R. (1996). Die mentale Repräsentation von Konzepten, Wörtern und Figuren. In Grabowski, J., Harras, G. & Herrmann, Th. (Hrsg.), *Bedeutung - Konzepte, Bedeutungskonzepte*. Opladen: Westdeutscher Verlag.
- Hoffhenke, M., Wachsmuth I. (1997). *Dynamische Konzeptualisierung mit imaginalen Prototypen*. SFB 360 Report 97/9, Universität Bielefeld.
- Hummel, J. E. (1998). Where view-based theories break down: The role of structure in shape perception and object recognition. In: Dietrich, E. & Markman, A. (Eds.). *Cognitive dynamics: Conceptual change in humans and machines*. MIT Press, Cambridge, MA.
- Jung, B. (1997). *Wissensverarbeitung für Montageaufgaben in virtuellen und realen Umgebungen*. Dissertationen zur Künstlichen Intelligenz, 157, infix, Sankt Augustin, Germany.
- Jung, B., Hoffhenke, M., Wachsmuth, I. (1998). Virtual assembly with construction kits. *Proceedings of the 1998 ASME Design for Engineering Technical Conferences (DECT-DFM '98)*.
- Kessler, K. & Rickheit, G. (im Druck). Eine Netzwerkkonzeption von Konzeptdynamik: Begriffsklärung, Modellvorstellungen und experimentelle Ergebnisse. Eingereicht bei *Kognitionswissenschaft*.
- Klix, F. (1980). Die allgemeine psychologie und die Erforschung kognitiver Prozesse. *Zeitschrift für Psychologie*, 188, 117-139.
- Kosslyn, S. M. (1980). *Image and mind*. Harvard University Press, Cambridge (MA).
- Kosslyn, S. M. (1994). *Image and brain: the resolution of the imagery debate*. MIT Press, Cambridge (MA).
- Marr, D. & Nishihara, H. K. (1978). Representation and recognition of the spatial organization of three-dimensional shapes. In *Proceedings of the Royal Society of London B*, 200, 269 - 294.
- Miller G. A., (1956). The magical number seven, plus or minus two. Some limits on our capacity for processing information. *Psychological Review*, 63, 81-97.
- Paivio, A. (1979). *Imagery and verbal processes*. Hillsdale (N.J.): Lawrence Erlbaum Associates.
- Rickheit, G. & Strohner, H. (1993). *Grundlagen der Kognitiven Sprachverarbeitung*. Tübingen: Francke.
- Shastri, L. & Ajjanagadde, V. (1990). *From simple associations to systematic reasoning: a connectionist representation of rules, variables and dynamic bindings*. (Technical report MS-CIS-90-05). Department of Computer Science, University of Pennsylvania.
- Siddiqi, K., Shokoufandeh, A., Dickinson, S. J. & Zucker, S. W. (1998). Shock graphs and shape matching. *Proceedings of the Sixth International Conference on Computer Vision*, Bombay, India.
- Smith, E. E., Shoben, E. J. & Rips, L. J. (1974). Structure and process in semantic memory. *Psychological Review*, 81, 214-241.
- Smolensky, P. (1987). *A method of connectionist variable binding*. (Technical Report CU-CS-356-87), Boulder: University of Colorado, Department of Computer Science.
- Tarr, M. J. & Bühlhoff, H. H. (1998). Image-based object recognition in man, monkey and machine. In *Cognition*, 67, 1-20.
- Treisman, A. & Schmidt, H. (1982). Illusory conjunctions in the perception of objects. *Cognitive Psychology*, 14, 107-141.
- Wachsmuth, I., Jung, B. (1996). Dynamic conceptualization in a mechanical-object assembly environment, *Artificial Intelligence Review*, 10 (3-4), 345-368. Reprinted in Mc Kevitt, P. (Ed.) (1996). *Integration of natural language and vision processing (Vol. IV)*, 191-214, Dordrecht: Kluwer.