Everything you always wanted to know about

English relative clauses

. . . but were afraid to ask

What is a relative clause?

The phenomenon

A relative clause (RC) is a kind of subordinate clause that modifies a noun or nominal.

i. Micheal hates NP[**people** [who listen to Robbie Williams]]

What is a relative clause?

The phenomenon

RCs are usually introduced by an **element R** that is **anaphorically related** to the modified nominal.
This element derives its semantic interpretation from its antecedent, the external **head** of the relative clause.

i. Micheal is [**the guy** [**who** stole my car]].

HEAD R
antecedent

What is a relative clause?

The phenomenon

RCs need an element R that stands in an anaphoric relationship to the head, but R need not be a pronoun.

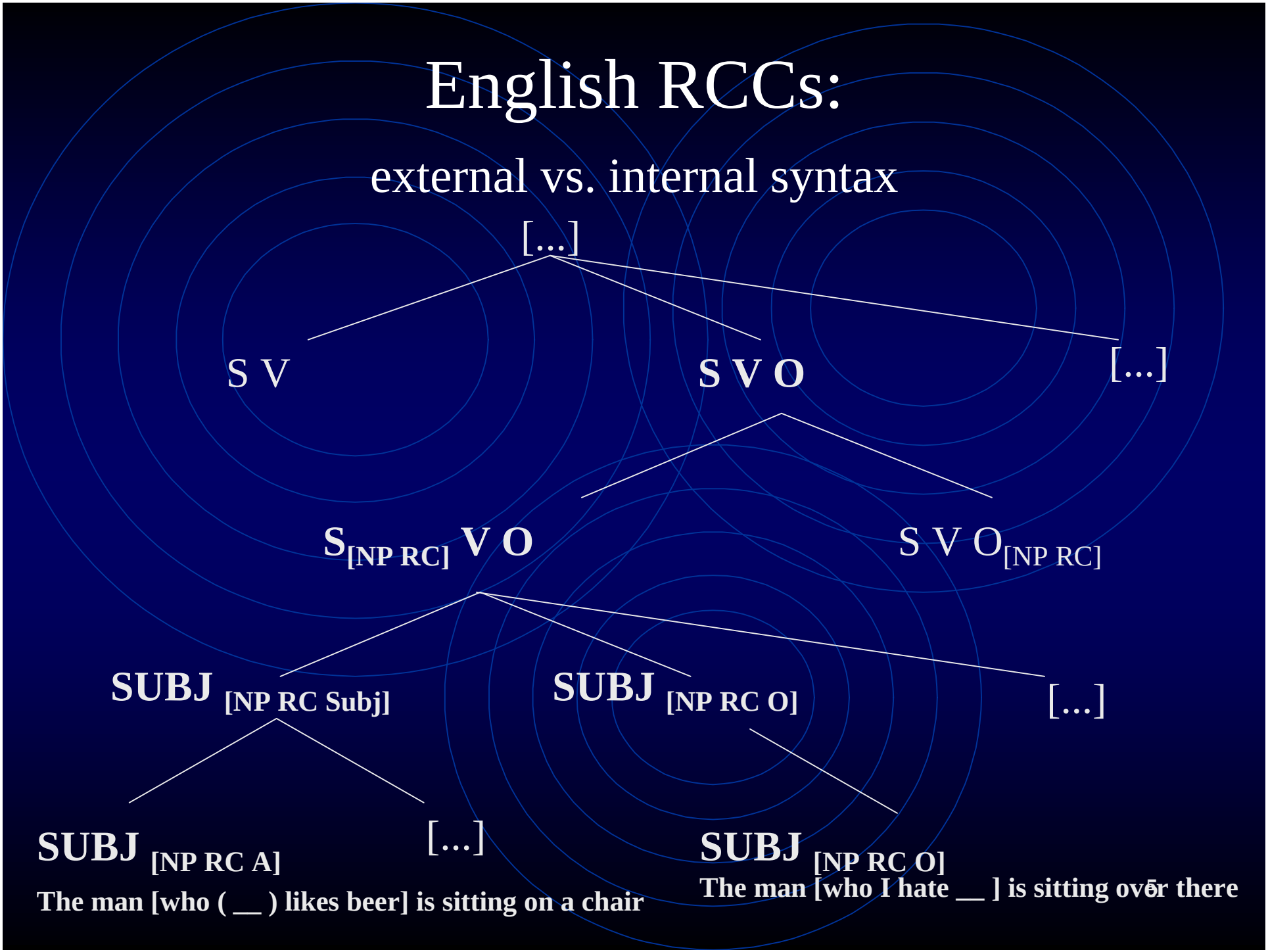
i. The guy **who** you are dating ___ is an idiot. (WH relative)
ii. The film **that** you like ___ is stupid. (*that* relative)
iii. The music ∅ you like ___ is characterless. (bare relative)

canonical position of argument of transitive verb

Terminology:
psycholinguistics: ***gap*** (R is ***filler***)
linguistics: trace (R_{prn} is left-dislocated wh-element)

English RCCs:

external vs. internal syntax



Some more types:
internal syntax

R, the relativised element, can have a range of functions:
(**pronominal; wh-relatives**)

i.	my friends [**who** (__) like MJ]	A (Subj RCtrans)
ii.	my friends [**who** (__) suck]	S (Subj RCintrans)
iii.	her boyfriend [**who(m)** I hate__]	Object
iv.	the books [**which** I referred to __]	Comp Prep
v.	the day [**when** you were born __]	Adjunct Time
vi.	a place [**where** you can relax __]	Adjunct Place
vii.	the reason [**why** you are here]	Adjunct Reason

I ignore relativization on genitive NPs within a subject|object NP etc. for the moment

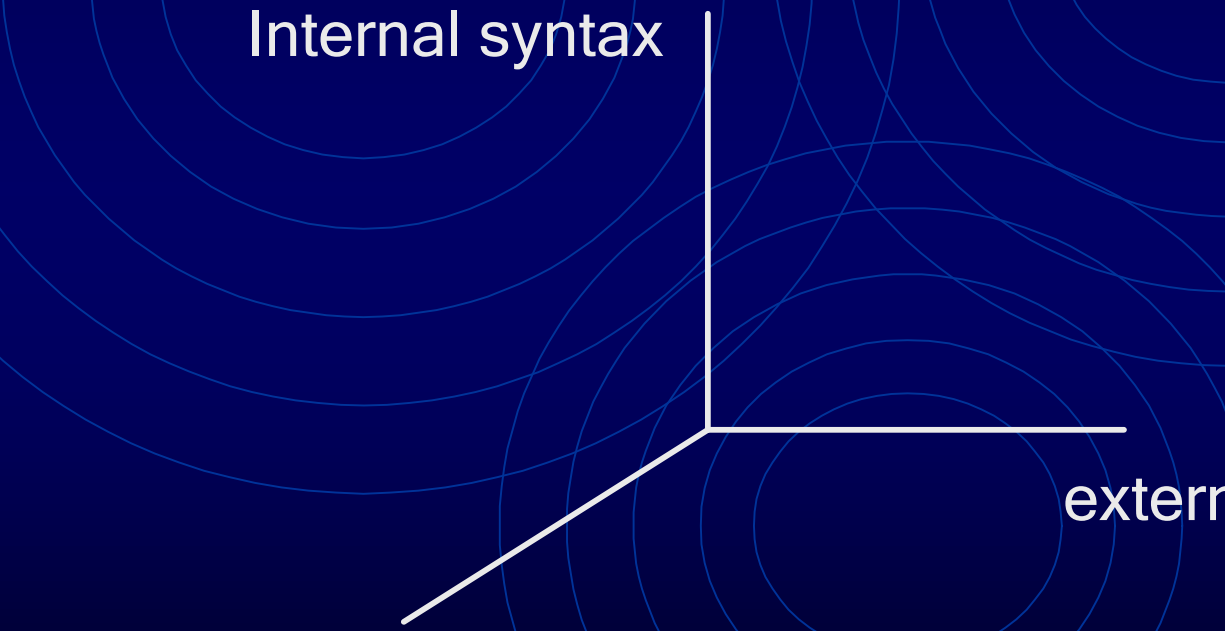
Some more types:
internal syntax

R, the relativised element, can have a range of functions:
(**non-pronominal; *that* and bare-relatives**)

i.	some guys [that __ saw me]	A (Subj RCtrans)
ii.	some guys [that ___ suck]	S (Subj RCintrans)
iii.	a key [(zero) she found ___]	Object
iv.	the books [(zero) I referred to ___]	Comp Prep
v.	the day [(zero) you were born ___]	Adjunct Time
vi.	a place [(zero) you can relax ___]	Adjunct Place
vii.	the reason [(zero) you got angry ___]	Adjunct Reason

I ignore relativization on genitive NPs within a subject|object NP etc. for the moment

so, the phenomenon is rather complex...



Interlude:

"Journey to the center of Grammar"

And … what is explanation?

i. Alex likes $_{NP}$[the woman $_{S}$[who wears **hats]]**.
ii. ***What**$_i$ does Alex like the woman who wears $\_\_{}_i$?
iii. Alex likes the woman who wears **what**?

Explanations* in formalist approaches:

⇒ identify|discover the principle of UG which governs the phenomenon:
e.g. relative clauses are 'barriers' (Chomsky 1986)

But for many linguists this is
not exactly a very revealing explanation

Different explanations:
core assumptions

Following usage-based approaches to grammar, I assume that...

I. There is a profound ***correspondence*** between ***competence*** and ***perfomance***

II. Grammars have ***conventionalised syntactic structures*** in proportion to their ***degrees of preference***

III. Degrees of preference are (to a large extent) determined by ***processing cost***

IV. Degrees of preference can be assessed by means of ***quantitative corpus-linguistic investigation***

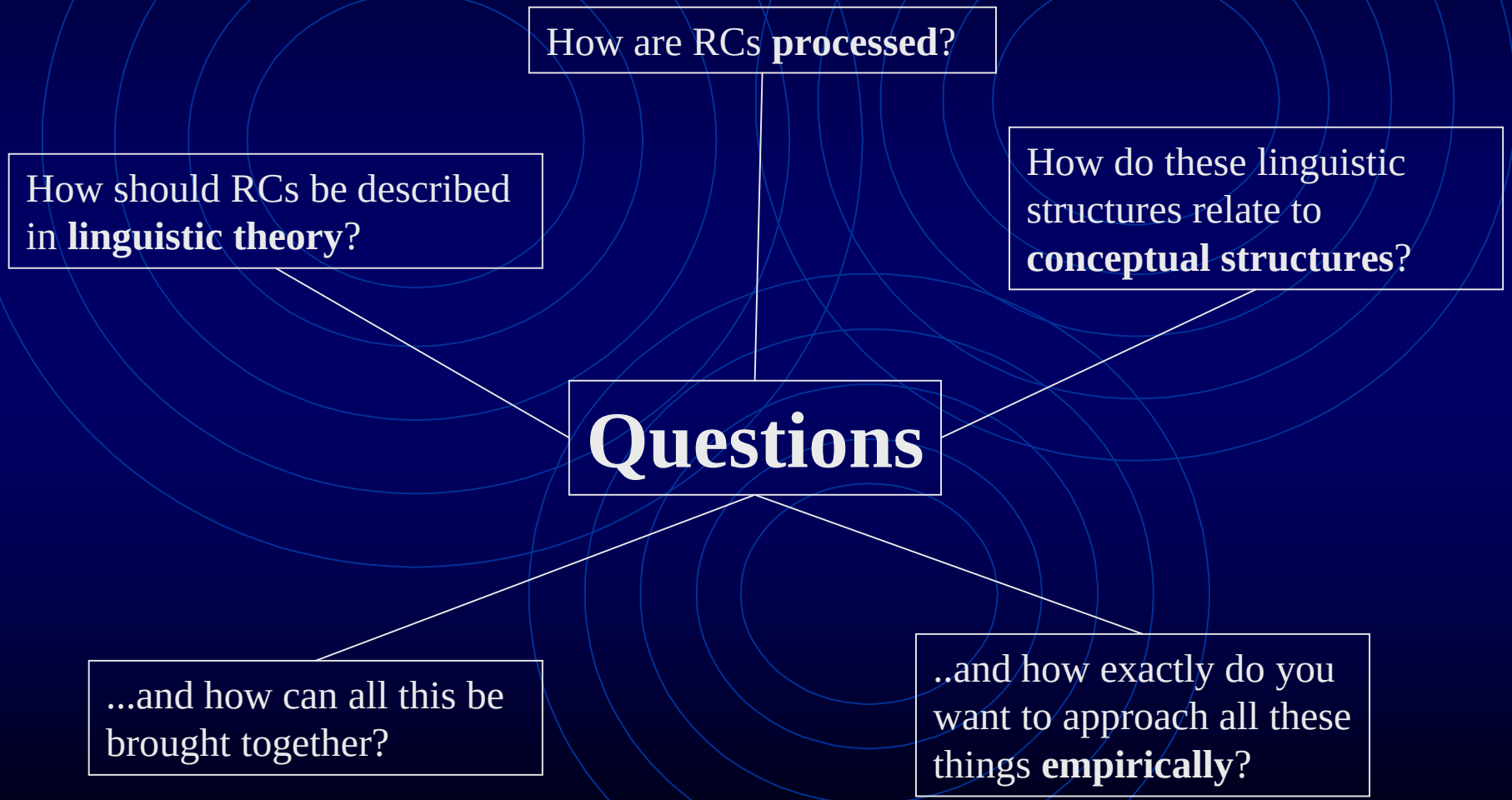
V. Degrees of preferences of linguistic structures are reflections of ***preferred instructions*** to the ***conceptual processing system*** on how to conceptualize, i.e. mentally simulate, the decribed scene

∴ **We need to know *how the conceptual system operates* to understand language processing**
∴ **Most importantly, we need to know what the respective *representational formats* are (i.e. amodal? - multi-modal?)**

Different explanations: core assumptions

By implication, all these hypotheses apply to relative clauses.

Things to know about relative clauses



Things to know about relative clauses

How should RCs be described in linguistic theory?

- are there coherent sub-classes of RC constructions that need to be identified?
- if so, how are these constructions related to each other
- and how are these constructions related to other - formally identical|functionally similar - constructions?

How do these linguistic structures relate to conceptual structures?

How are RCs processed?

Questions

...and how can all this be brought together?

..and how exactly do you want to approach all these things empirically?

Things to know about relative clauses

How are RCs processed?

- are certain types of RCs harder to process than others?
- if so, why?
- what are the properties of these types that make them hard to process
- are there certain sets of feature combinations that result in processing difficulty?

How should RCs be described in linguistic theory?

Questions

How do these linguistic structures relate to conceptual structures?

...and how can all this be brought together?

..and how exactly do you want to approach all these things empirically?

Things to know about relative clauses

How should RCs be described in linguistic theory?

How are RCs processed?

Questions

How do these linguistic structures relate to conceptual structures?

- what are the corresponding conceptual structures and how do these look like (representational formats)

...and how can all this be brought together?

..and how exactly do you want to approach all these things empirically?

Things to know about relative clauses

How should RCs be described in linguistic theory?

How do these linguistic structures relate to conceptual structures?

How are RCs processed?

Questions

...and how can all this be brought together?

..and how exactly do you want to approach all these things empirically?

>>Data, Parameters, Methods<<

>>thoughts on Analogical Modelling<<

Things to know about relative clauses

How are RCs processed?

How should RCs be described in linguistic theory?

How do these linguistic structures relate to conceptual structures?

Questions

..and how exactly do you want to approach all these things empirically?

...and how can all this be brought together**?**
'**Embodied Construction Grammar**'

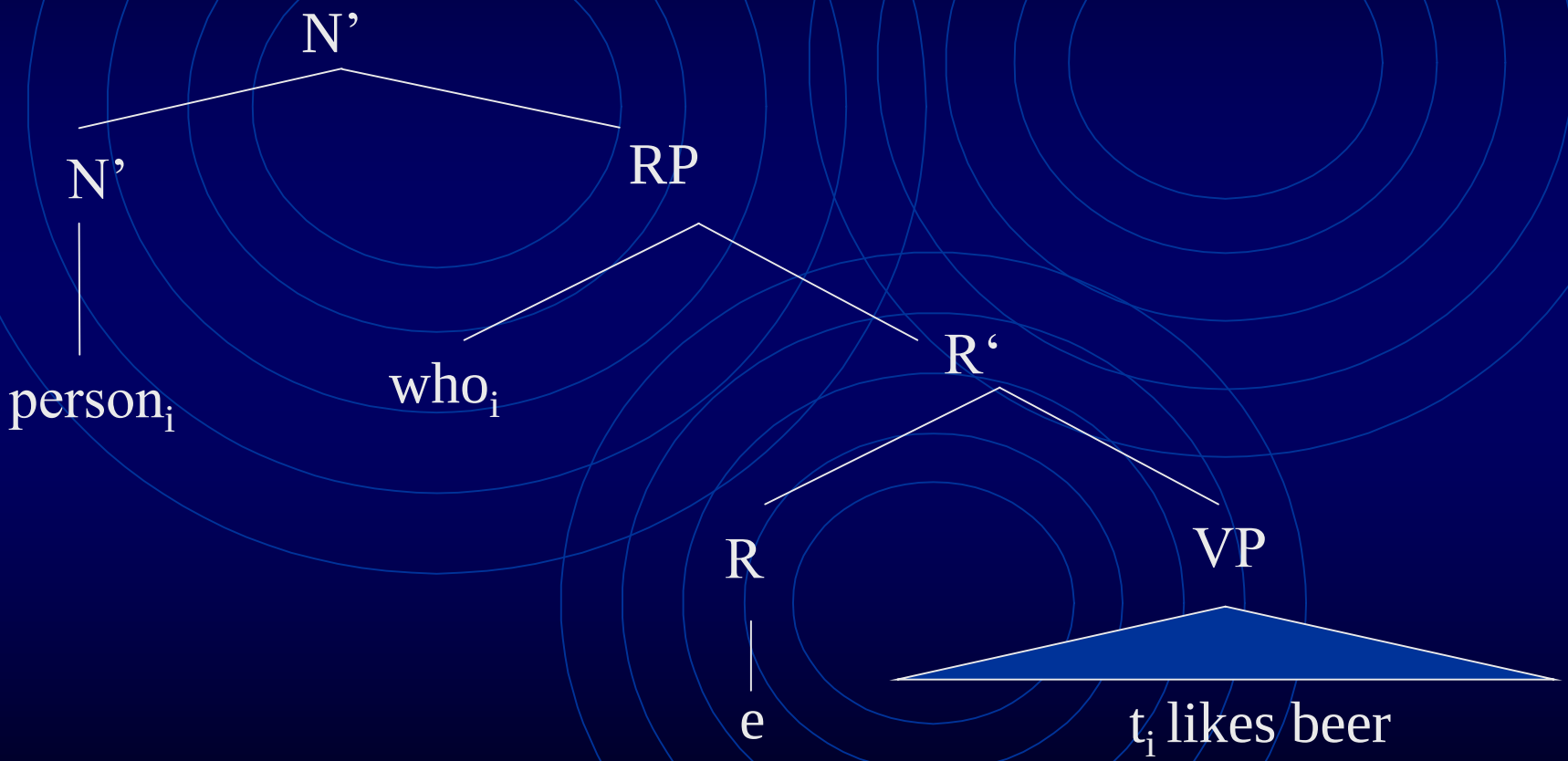
Theory I:
Linguistic theory
(RCs in sign-based unificational theories)

"Data-collecting leaves one ill-prepared for the complexities of theoretical thinking."

Esa Itkonen

back

Relative Clauses in Government and Binding Theory



phrase is projected from an invisible (phonetically empty) R element

…but I don't like theories that include unobservables…

Invisible elements are routinely assumed in modern syntactic discussions (X0 traces, PRO, NP-traces, wh-traces)…

…but have ***at best theory internal motivations*** (e.g. required by Projection Principle)

Recent versions of HPSG (and other constraint-based lexicalist frameworks) promise to ***do away with them***.

Head-Driven-Phrase-Structure-Grammar (HPSG)

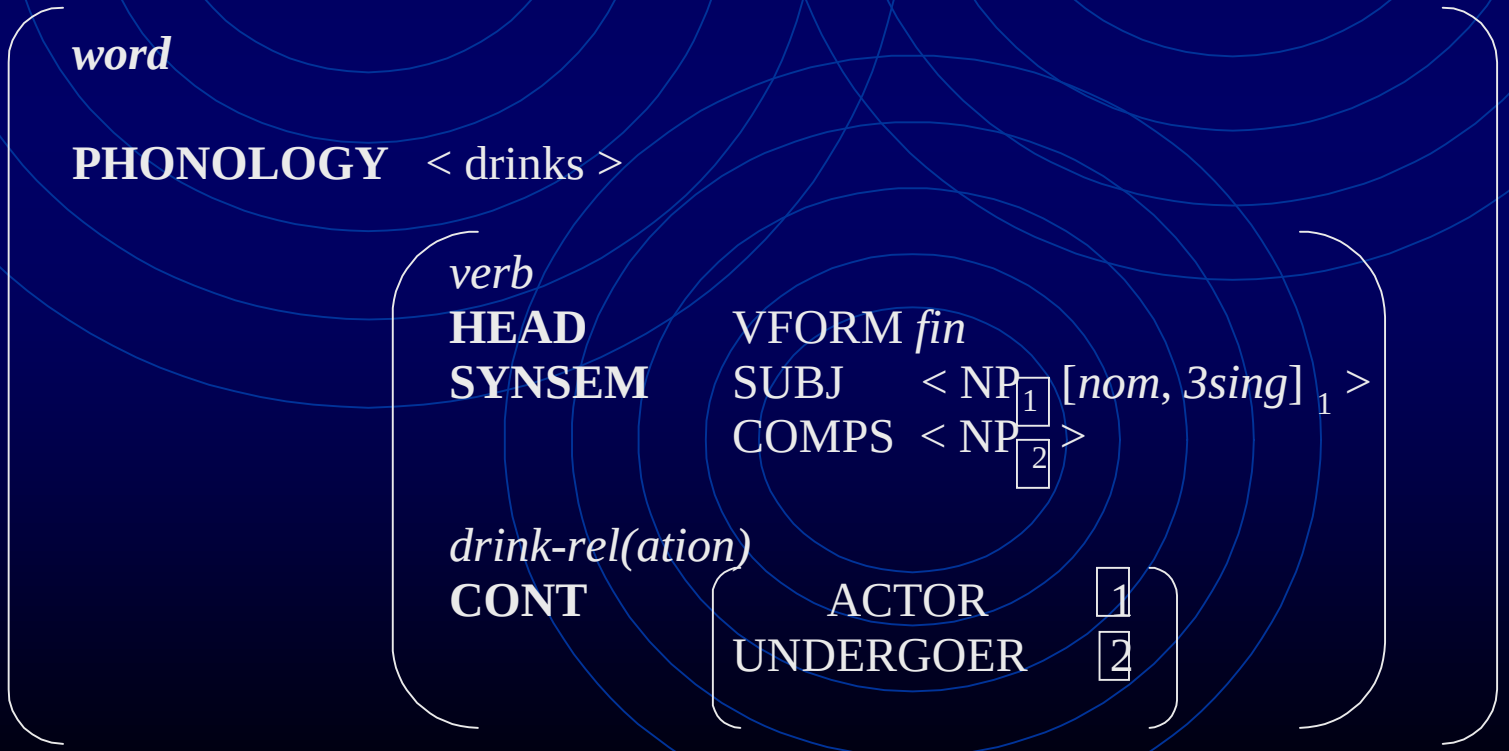
Some basics

- fundamental utterance type recognized in HPSG is the *sign*
- two subtypes: *word* and *phrase*
  - so, just as lexical entries are descriptions of (constraints on) feature structures (FS) of type *word*, schemata are descriptions of FS of type *phrase*
- **the grammar of a language is then just the specification of its types and the constraints that govern those types**

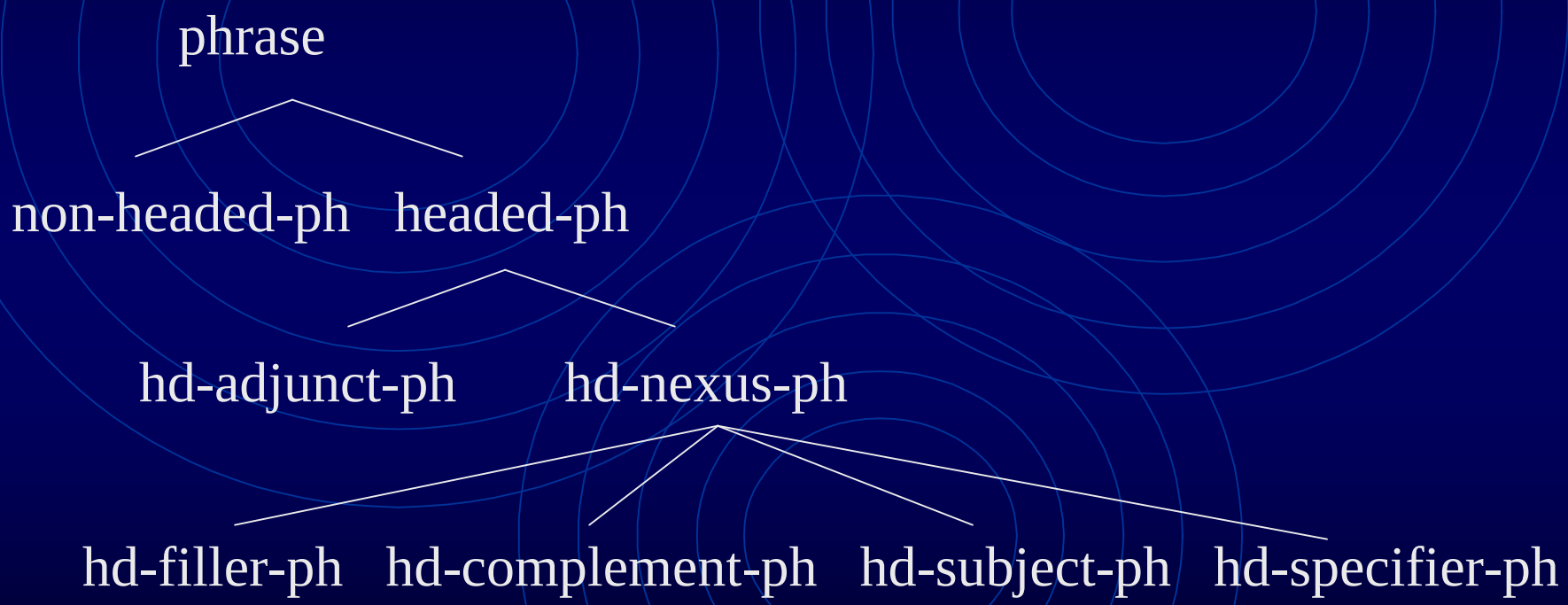
Description of signs (types):

Feature stuctures (FS)

Attribute Value Matrices (AVM) are employed as a **description language of FS**: Example *drinks*



types hierarchies: an example

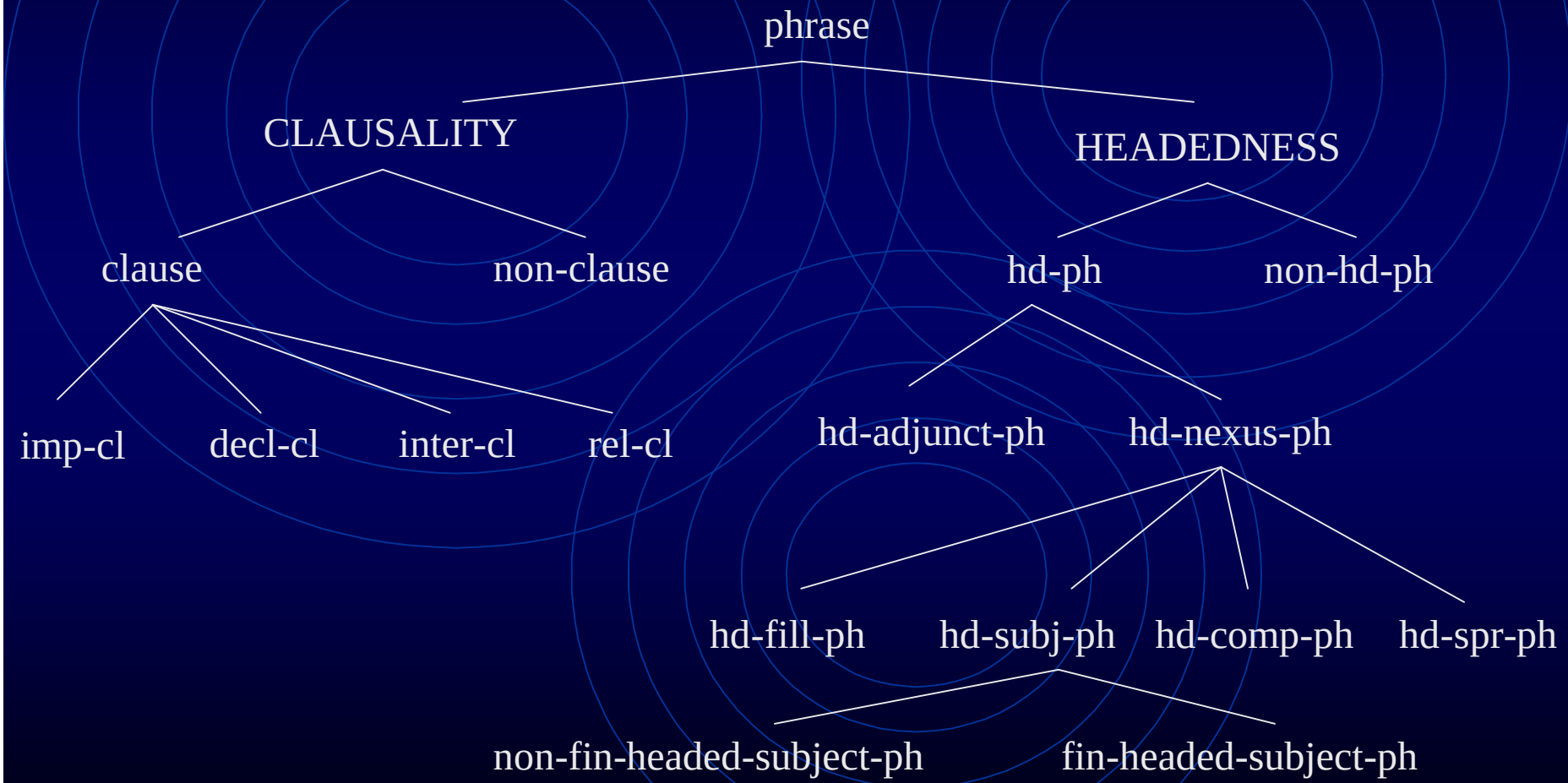


Coming back to relative clauses…

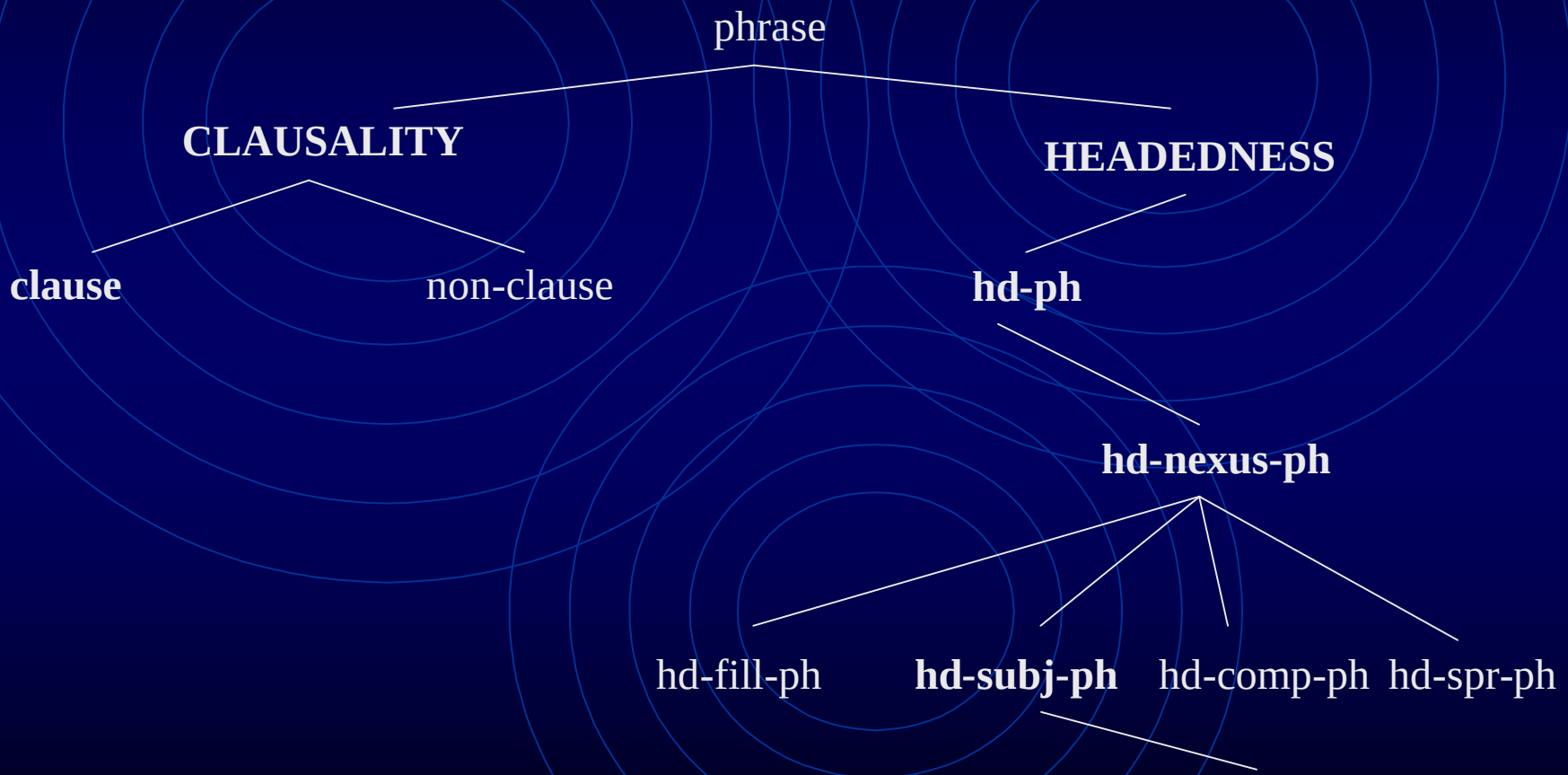
These clauses share the **same phrase structure**:

(i)	**{who|whose flat-mate} usually drinks a lot of milk**	**relative**
(ii)	Stephen usually drinks a lot of milk.	declarative
(iii)	{Who|Whose flat-mate} usually drinks a lot of milk?	interrogative

phrase types:
headedness and clausality



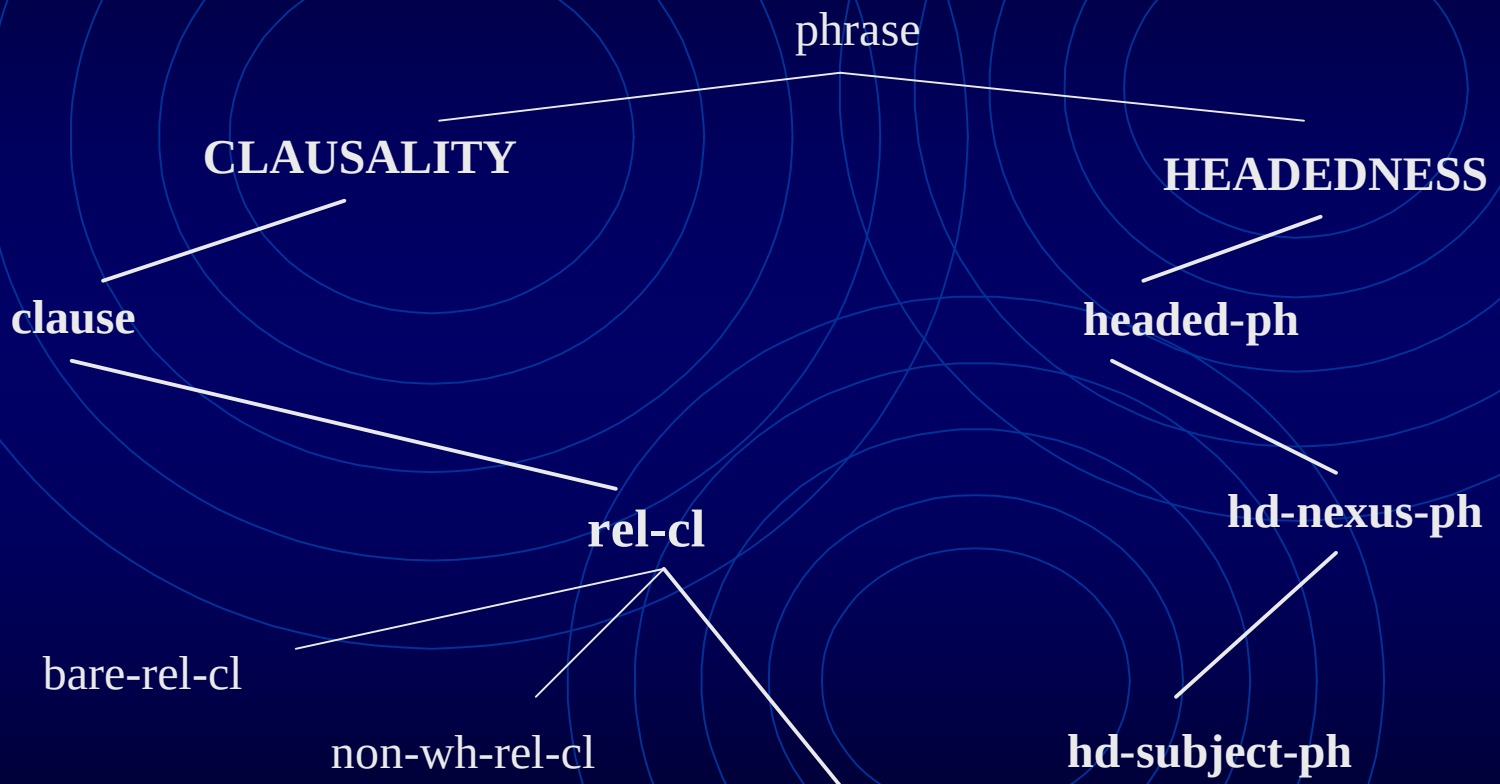
phrase types:
formal generalisation



(i) - (iii) are all instances of a common type: clause

fin-headed-subject-ph

phrase types:
further specification



…and (i), who usually drinks a lot of milk, can be specified as:

Good thing we have a sign-based theory

semantic differentiation

TYPE	**CONSTRAINTS**		**ISA**
decl-cl	[CONTENT	*assertion**]	clause
inter-cl	[CONTENT	*question*]	clause
imp-cl	[CONTENT	*directive*]	clause

rel-cl

HEAD **MC -**
INV -
MOD [HEAD noun]

CONTENT *proposition*

But maybe we can even do more:

fine-grained functional differentiation

Relative clauses occur in different environments (**external syntax**) and show recurring lexico-grammatical patterns (**internal syntax and lexical material in slots**)

Maybe these instantiate different constructions with different functions?

Some functions of relative clause constructions

Referential RCs:

"Politicians who make big promises aren't trusted."

Semantics:

- information expressed is presented as an integral part of larger message (**restricts set of potential referents**, i.e. denotation)

Some functions of relative clause constructions

Attributive RCs:

"Politicians, who (by the way) make big promises, aren't trusted."

Semantics:

- information expressed is **supplementary**, additional
- does not restrict denotation
  - » Intonation: spoken as separate intonation unit
  - » But it is unclear whether speakers do actually reliably produce such cues.

Allbritton, D., McKoon, G. and Ratcliff, R. (1996) The reliability of prosodic cues for resolving syntactic ambiguity *J. Exp. Psychol. Learn. Mem. Cognit.* 22, 714–735

Some functions of relative clause constructions

Presentational relatives

“This is the woman, who I wanted to meet my whole life.”

Semantics:

- may be considered **focus-construction**; “mono-propositional”
- may or may not restrict denotation

Some functions of relative clause constructions

Infinitival *to* relatives:

"Things to do when you are on holidays."

Semantics:

- used in **ad hoc concept-formation**
- highly restrictive

Summary:
Guiding questions

1. Exactly how many (and which) RCCs should we assume as conventional parts of grammar?

Example:Adjuncts may appear in a number of positions

i. the book he **hastily** gave to Mary
ii. ?the book he gave **hastily** to Mary
iii. the book he gave to Mary **hastily**
iv. daß **schnell** jemand den Teich leer fischt
v. daß jemand **schnell** den Teich leer fischt
vi. daß jemand den Teich **schnell** leer fischt
(Müller 2006)

Are these examples of different constructions?

In GB, one can turn to movement operations
⇒ Adjunct serialization can be explained by scrambling;
But such mechanisms are not availble in nonderivational theories
∴ We stipulate more and more constructions

Questions

1.1 How do they relate to each other?

- (formally and functionally)

1.2 And how do they relate to non-RC-constructions

- (e.g. other focus constructions)?

1.3 Can we actually draw clear-cut distinctions between constructions?

- The conference planned by non-governmental organisations was about globalisation (reduced relative?)

back

Theory II: The experiential view of language comprehension

"A major function of language is thus to enable us to experience the world at proxy"

Philip N. Johnson-Laird (1983)

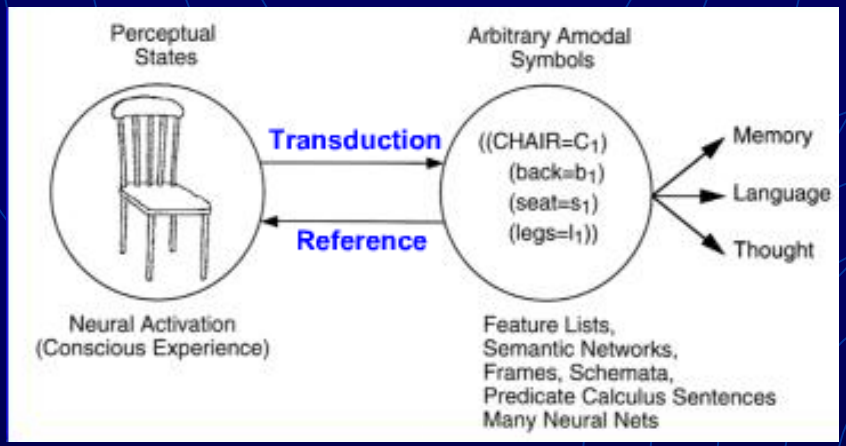
The experiential view in a nutshell

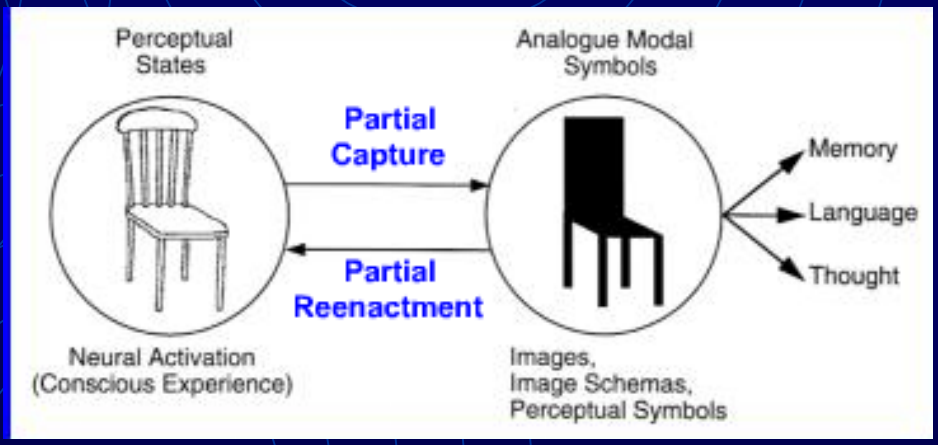
- All mental representations are experiential, i.e., related to perception and action.
  - referent representations (RR)
  - linguistic representations (LR)
- RR are traces laid down in memory because of perceptions of and interactions with the environment.
  - RR are multi-modal
  - RR are schematic → because of attentional limitations
- LR are laid down, as linguistic information is being received or produced.
- All (RR & LR) constructions are interconnected
- LR are also connected to RR

- How are these interconnections established?
- The main mechanism is co-occurrence (e.g., Hebb, 1949)
- Certain entities/events in the environment tend to co-occur.
  - Because of these spatio-temporal co-occurrences, combinations of entities and events become part of the same experiential trace.
  - Because they co-occur with the entities/events, linguistic become associated with RR

The experiential view of language comprehension

amodal modular vs **modal non-modular**
views of concepts





Figures 1 & 2 (from: Barsalou 1999)

Why has the the perceptually based view not enjoyed acceptance?

Actually, it has been dominating human thinking...

- Aristotle & Epicurus (4th century BC): Representations that underly cognition are imagistic
- British Empiricists: Locke (1690), Berkeley (1710), Hume (1739) viewed cognition in this manner
- Even later nativists (Kant 1787, Reid 1764) had images play a central role in their theories
- And even recent philosophers (Russell 1919) have incorporated images into their theories

Why has the the perceptually based view not enjoyed acceptance?

Then came ordinary language philosophy and behaviorism...

- Mental states (MS) were banished from consideration in much of the scientific community
- MS are unscientific and lead to confused views of human nature (cf. Ryle 1949, Watson 1913, Wittgenstein 1953)
- Attacks on mentalism included a critique of images, so when the former went down the drain, so did the latter.

Daniel:

The goal of these attacks was not to exclude images from mentalism but to eliminate mentalism altogether (as a result, image-based theories of cognition disappeared with theories of cognition)

Why has the the perceptually based view not enjoyed acceptance?

The birth of ‘Cognitive Science’: A meeting at MIT on September 11, 1956 the symposium on information theory.

- **Shannon** had formulated the idea of information engaged in processing information and had talked about coding theory
- **Newell** and **Simon** had presented their logic theory machine
- **Rochester** and colleagues presented a computer implementation of Hebbs neurophysiological theory of cell assemblies
- **Chomsky** talked about his idea that linguistic knowledge involves rewrite rules and transformation
- **Miller** presented his classic paper about the magical number 7

Why has the the perceptually based view not enjoyed acceptance?

In conclusion:

The mind is an information processing system

Why has the the perceptually based view not enjoyed acceptance?

‘Cognitive Science’ brought us…

- new representational languages (feature lists, frames, schemata, semantic nets, production systems, connectionism)

- these differed from earlier ones in their relation to perception (cf. figures 1&2)

Daniel:

Early theories: assumed that cognitive representations (thought) utilize perceptual representations (FIG 1)

Why has the the perceptually based view not enjoyed acceptance?

amodal modular vs **modal non-modular** views of concepts

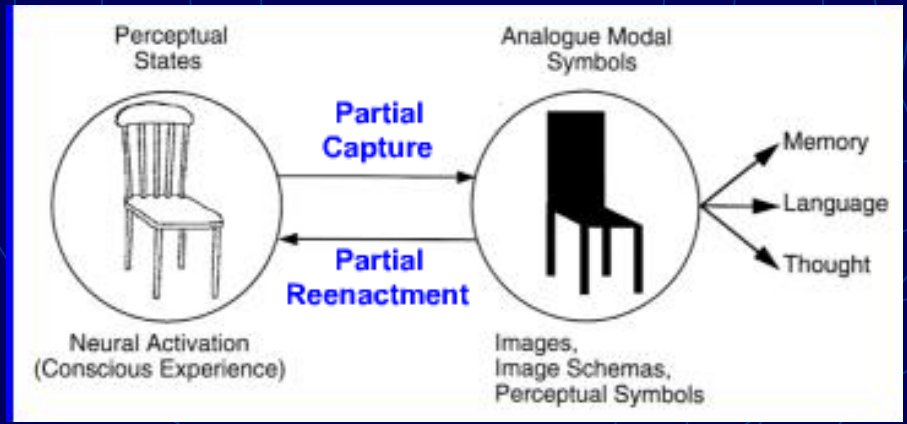


Figure 1: system extracts **a subset of a perceptual state** and stores it for later use as a **symbol**

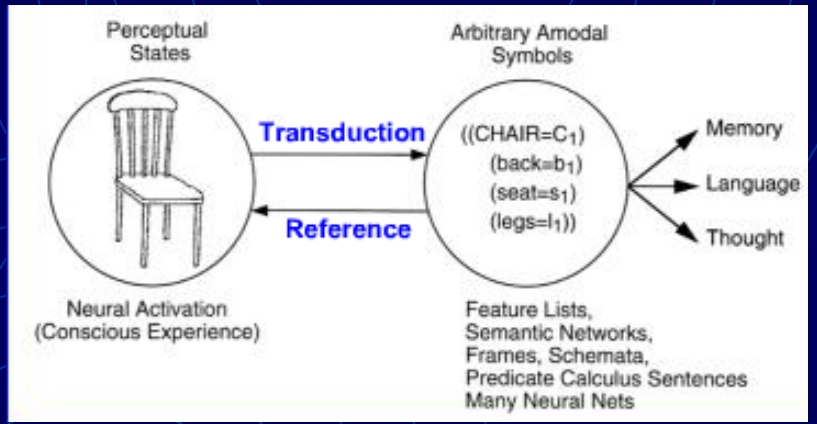
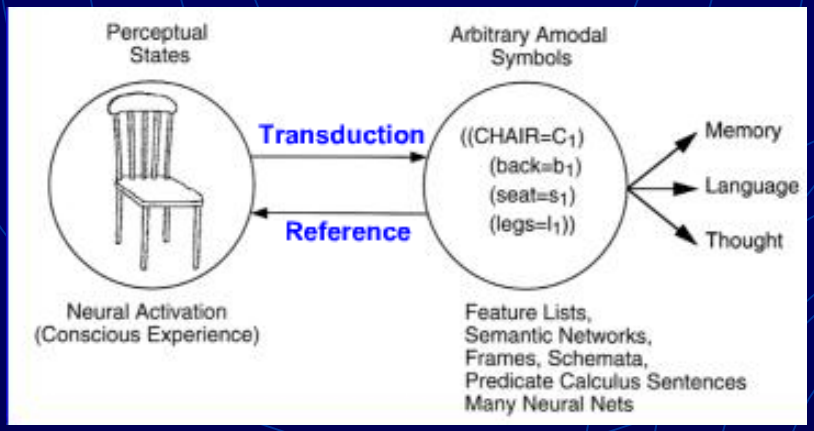


Figure 2: an **amodal symbol system transduces** a subset of a perceptual state **into a completely new representation language** that is inherently **nonperceptual**

Why has the the perceptually based view not enjoyed acceptance?



Symbols in these systems are **amodal** and **arbitrary**

- They are amodal because their internal structures bear **no correspondence to the perceptual states** that produced them
  - Because of their amodal character, these symbols are arbitrary

Daniel:

Words in L do not literally constitute the content of these representations, it is assumed that close amodal concepts do

s the the perceptually based view not enjoyed acceptance?

mbols bear an important relation to words in a language:

Theorists of these accounts typically use linguistic forms to represent amodal symbols as in:

Feature lists:
CHAIR
seat
back
legs

Schemata, frames, predicate calculus:
EAT
AGENT = horse
OBJECT = hay

Why has the the perceptually based view not enjoyed acceptance?

Symbolic **thought** is assumed to be analogous in many important ways to **language** (also sharing design features like ‘systematicity’, ‘recursivity’):

language processing involves the sequential processing of amodal symbols

conceptual processing involves the sequential processing of amodal symbols

(cf. Fodor & Pylyshyn 1988)

Aside: Many connectionist models (e.g. McClelland et al. 1986 and Rumelhart et al. 1986), too, put emphasis on such amodal symbols; here, connectionism is a means of implementation and no theoretical alternative

Why has the the perceptually based view not enjoyed acceptance?

Strengths:

- Capable of **representing types and tokens**
- Capable of **producing categorical inferences**
- Capable of **combining symbols productively**
- Capable of **representing propositions**
- Capable of **representing abstract concepts**

Amodal symbol system have made clear what any complete theory of cognition must account for!

So, what's the problem with amodal symbol systems then?

Problems:

- little empirical evidence that amodal symbols exist
  (cf. Glaser 1992, Seifert 1997)
- findings in neuroscience challenge amodal symbols, i.e. much cotegorical knowledge is grounded in sensory-motor regions of the brain
  (cf. Damasio 1989, Pulvermüller 1999)
- difficult to represent spatio-temporal knowledge
  (cf. Clark 1997)
- failure to provide a satisfactory of the transduction process
  (cf. Harnad 1987)
- symbol grounding problem (converse of transduction process)
  (cf. Searle 1980)
- Amodal symbol systems are too powerful
  (cf. Anderson 1978)

Daniel:

Damage to a particular sensory-motor region disrupts conceptual processing

Daniel:

As PoS often note: scientific theories must be evaluated on: falsifiability, parsimony, the ability to produce provocative hypotheses, the existence of direct evidence for their constructs, freedom of conceptual

Processing relative clauses

"[G]rammatical variation is highly correlated with processing ease […]"

Hawkins 2004:170

back

Issues in processing

<u>Psycholinguistic accounts:</u>

Some accounts assume an **association** is put up between **filler** and **gap** in on-line processing

This is the man$_i$ who$_i$ Peter likes __$_i$

filler gap

Terminology: ***filler/gap*** correspond to ***moved*** (left-dislocated) ***element*** and its ***trace*** in GB (and others)

The person$_i$ [who$_i$ you believe $_{(\_\_)}$ [that John saw $\_\_{}_i$]] is Mary

Theory III: Unifying linguistic theory and simulation based language understanding

Embodied Construction Grammar

(Bergen & Chang 2003)

back

Embodied construction Grammar in Simulation-Based Language Understanding (Bergen & Chang 2003)

- Understanding an utterance involves at least two distinct processes:
  - Analysis: to determine which construction the utterance instantiates
  - Simulation: according to the parameters specified by those constructions

back

A construction grammar formalism for simulation-based language understanding (Bergen & Chang 2003:9)

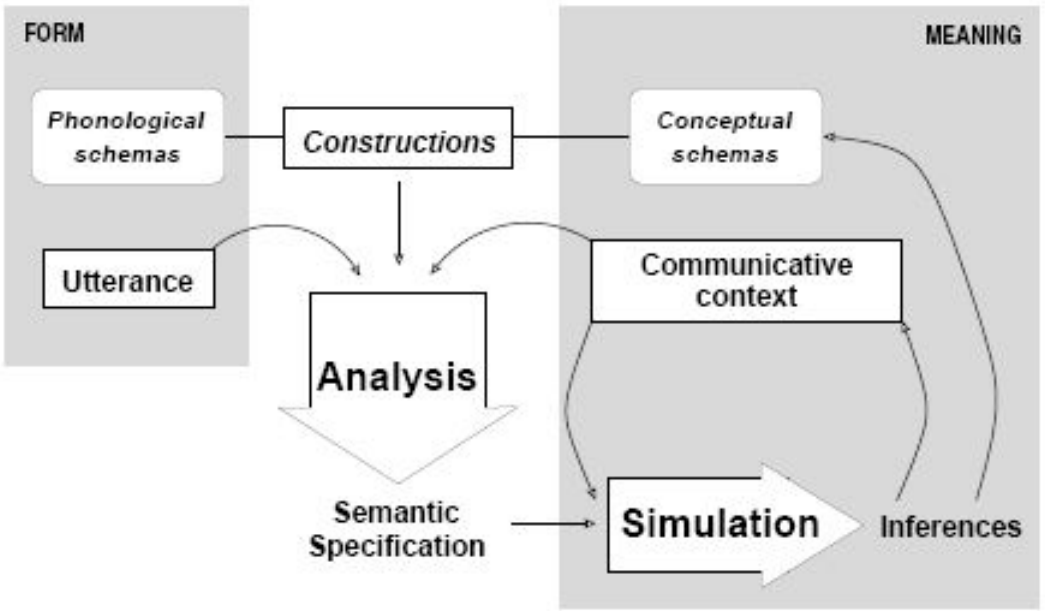
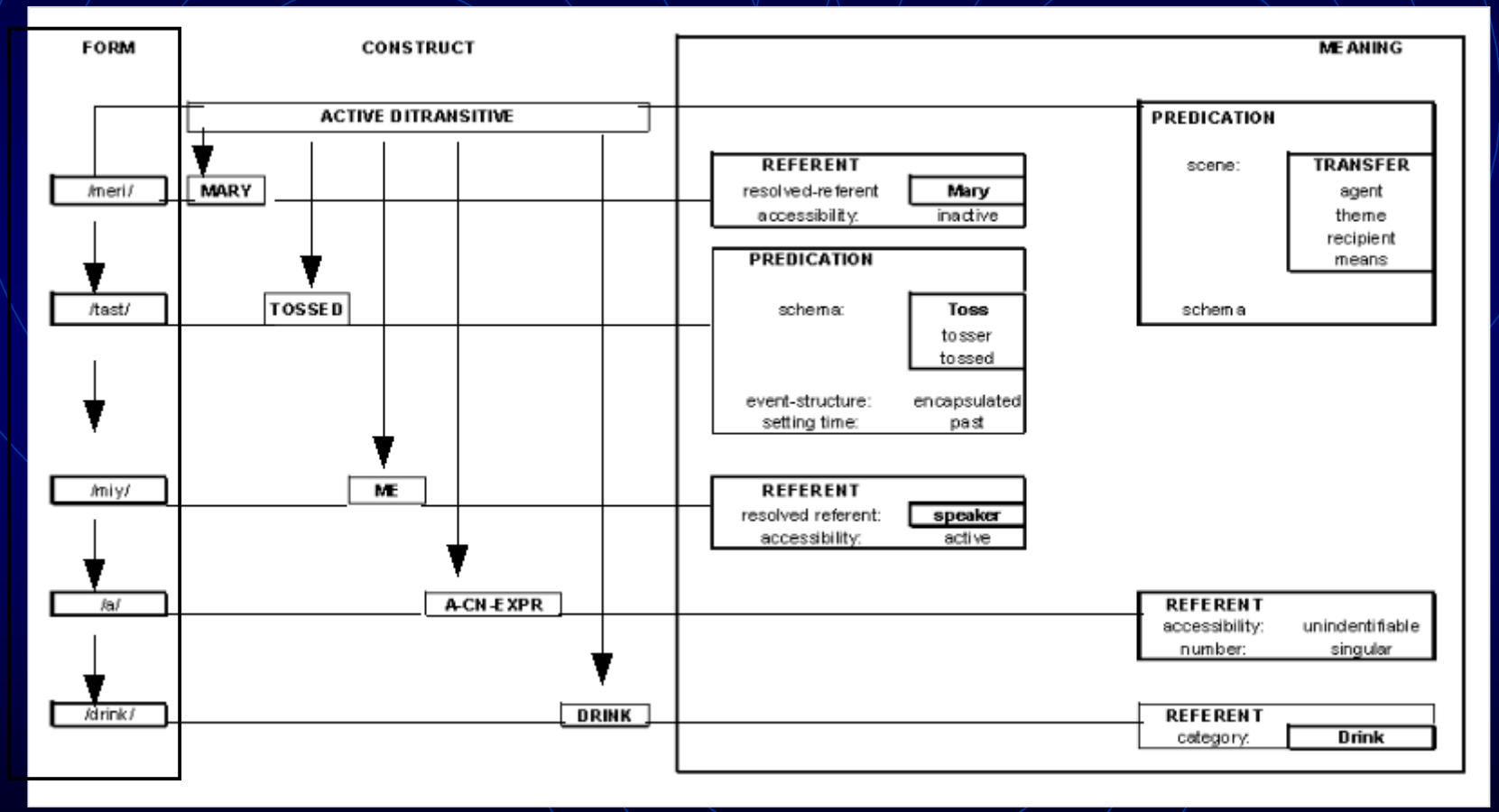


Figure 1: Overview of the simulation-based language understanding model, consisting of two primary processes: analysis and simulation. Constructions play a central role in this framework as the bridge between phonological and conceptual knowledge.

Construction formalism at work: licensing *Mary tossed me a drink*



ECG and language understanding

- dynamical processes that use the formalism:
  - **Construction analyzer** (a parser for form-meaning constructions)
    - must respect **phonological** and **conceptual constraints**
      - ≅ two interleaved procedures:
    1. **Search for candidate constructions** that may account for an utterance in context
    2. **Unification of structures** evoked by the constructions in a coherent ***SEMSPEC***

ECG and language understanding

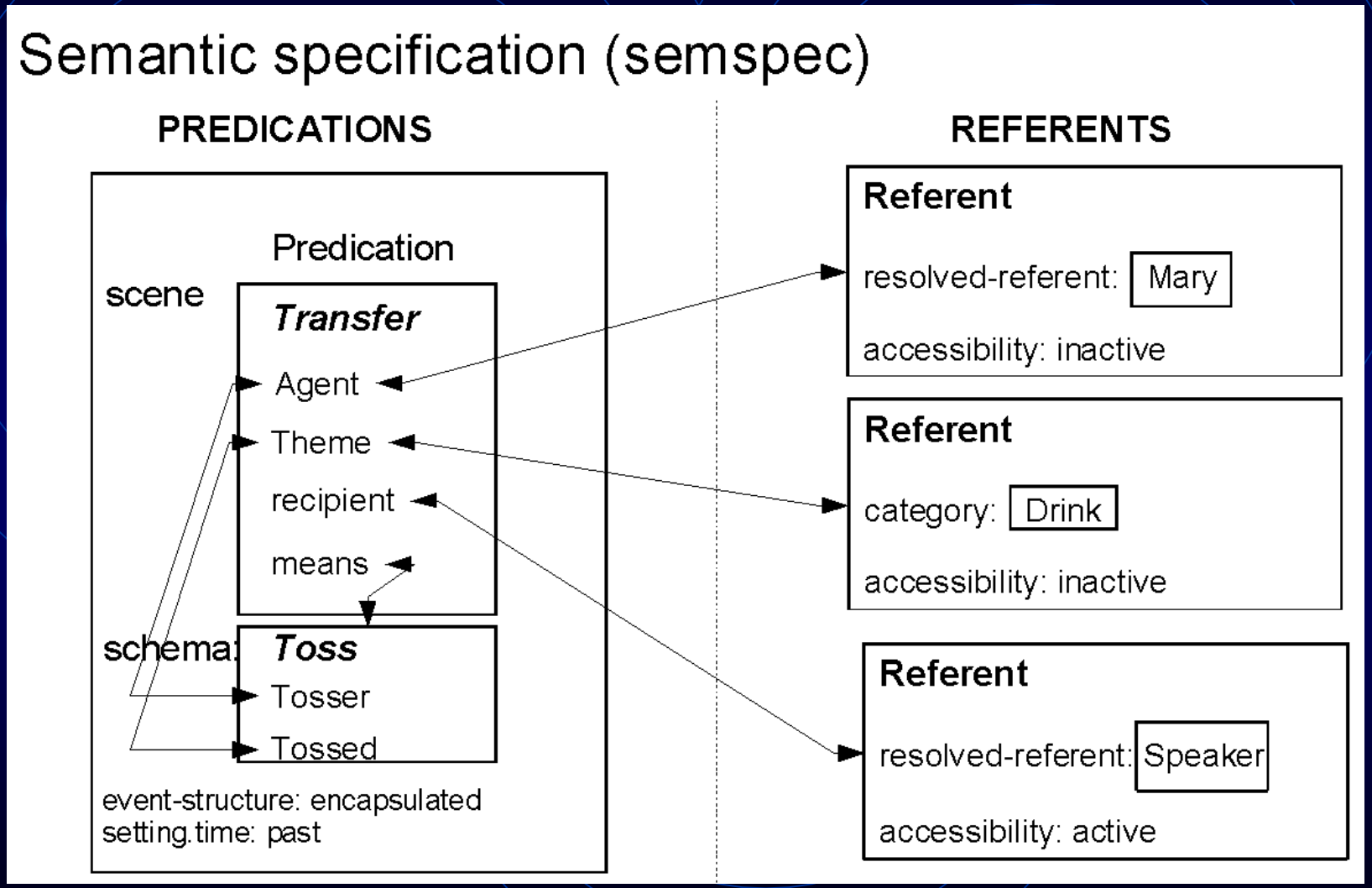
- Analysis begins with the phonological forms in an utterance triggering one or more constructions in which they are used (problem of ambiguity arises and disambiguation mechanism is necessary)
  - Activation of lexical constructions underlying the words *Mary*, *tossed*, *me*, *a*, *drink*

    → Various schemas evoked by the constructions are added to semspec
  - *E.g.* the indefinite article, *a*, cues the A-CN-EXPR construction (since *a* is part of its form pole)
  - Cued A-CN-EXPR construction has an additional com-noun slot to fill (satisfied by DRINK construct)
  - ACTIVE-DITRANSITIVE construct is triggered by by the other observed elements in the observed order and its constraints are then checked in context

Parsing

- Cued construction potentially supply top-down constraints on their constituents
  - Constructions and their constraints should be regarded not as deterministic but as fitting a given utterance and context to some quantifiable degree
    - Constructions and their constraints could be associated with probabilities or connection weights
      - cf. Narayanan and Jurafsky 1998
      - or in terms of association strength; Wiechmann (submitted)

ECG: SEMSPEC


Semantic specification (semspec)
PREDICATIONS
REFERENTS
Predication
scene
Transfer
Agent
Theme
recipient
means
schema
Toss
Tosser
Tossed
event-structure: encapsulated
setting.time: past
Referent
resolved-referent: Mary
accessibility: inactive
Referent
category: Drink
accessibility: inactive
Referent
resolved-referent: Speaker
accessibility: active

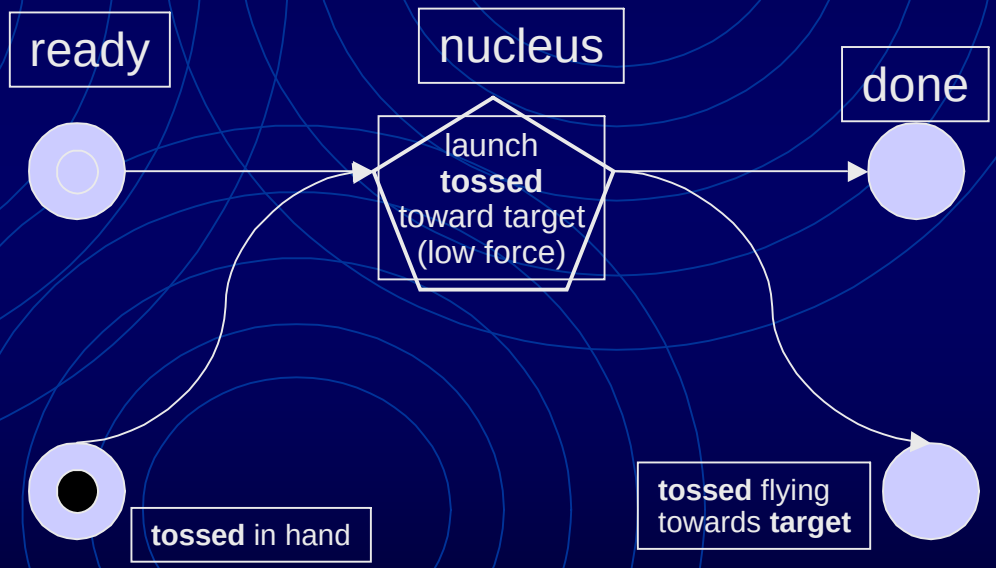
Simulative inference: execution schema (x-schema)

- **Executive schemas**, or **x-schemas**, are **dynamic representations** motivated in part by motor and perceptual systems, on the assumption that the same underlying representation used for executing and perceiving an action are brought to bear in understanding language about that action

- X-schema formalism is an extension of Petri Nets (Murata 1989) that can model sequential, concurrent, and asynchronous events
  - It also has natural ways of capturing features useful for describing actions (parametrization, hierarchical control, and the consumption and production of resources)

x-schemas

- Given the initial marking, the transition labeled 'nucleus' can fire, consuming tokens from each input place



x-schemas

- X-schemas model dynamic semantics by the flow of tokens
  - Tokens flow through the network along excitatory arcs according to th following rules:
    - When each of the transition's (excitatory) input places has a token, the transition is **enabled** and it can **fire**, consuming one token from each input place and producing one token in each output place
    - An x-schema execution corresponds to the sequence of markings that evolve as token flow through the net, starting from an initial marking

Tossing: percept vector

places

Places represent perceptual conditions or resources

tokens

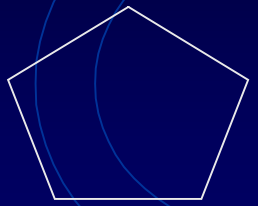
Tokens indicate that the condition is currently fulfilled
(2 in example, rest are **control states**

Overall state of x-schema is defined
as the distribution of tokens to places over the network
(this assignment is also called the **marking** of the x-schema)

arcs

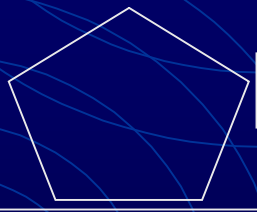
Directed arcs connect transitons to either **input places** (i.e. Places from which it has an incoming arc) or **output places**

most of the arcs are excitatory, but places and transitions can also be connected by **inhibitory** or **enabling arc**s



transitions

Transitions typically represent some action or change in condition of resources.
The ones shown each correspond to a complex action sequence with subordinate x-schemas (whose details are surpressed, or **encapsulated** at this level of granularity))

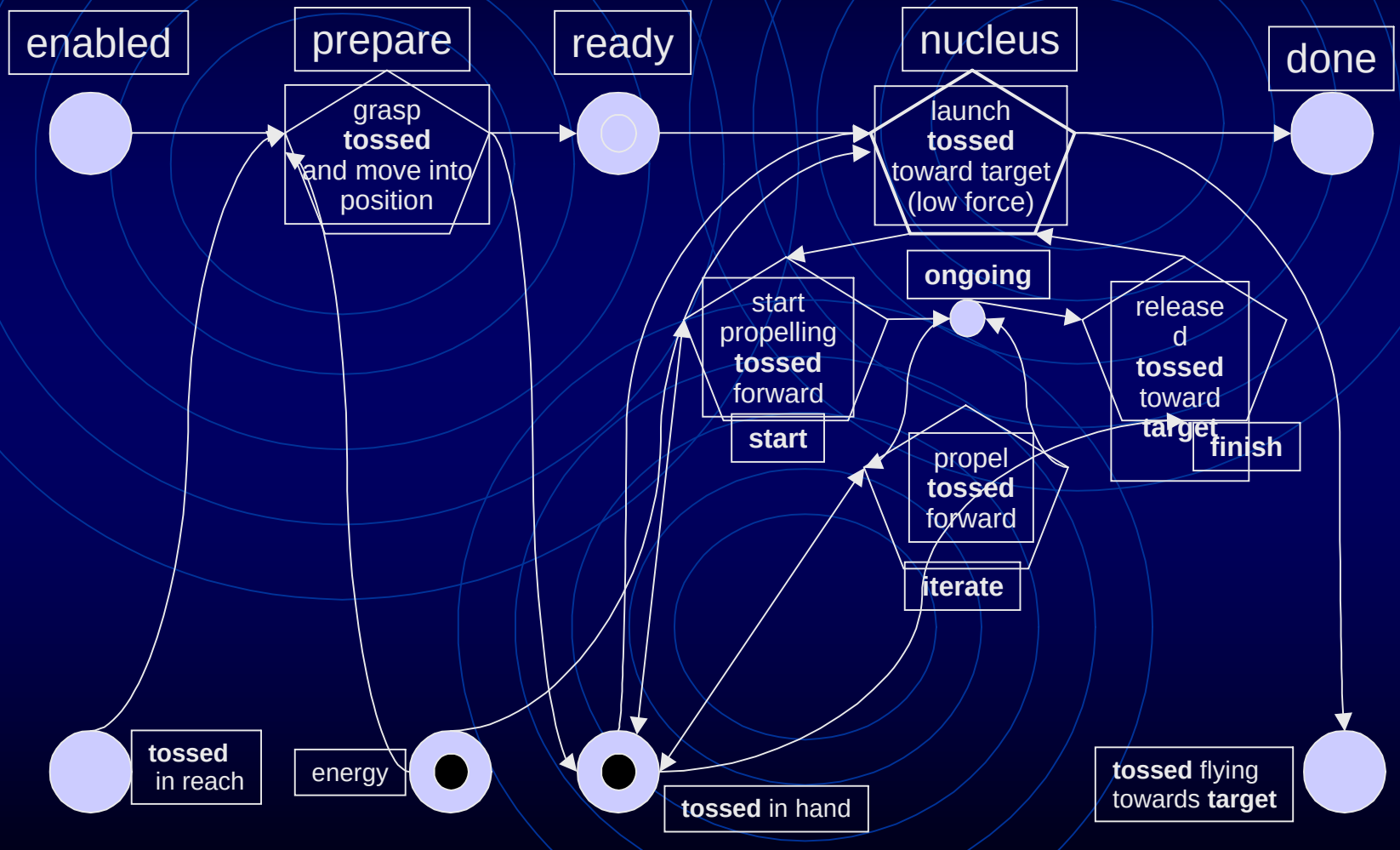


subordinate schema

NOTE that the transitions (and the place DW) also have labels relevant to the overall control of the action (cf. 3.2.2)

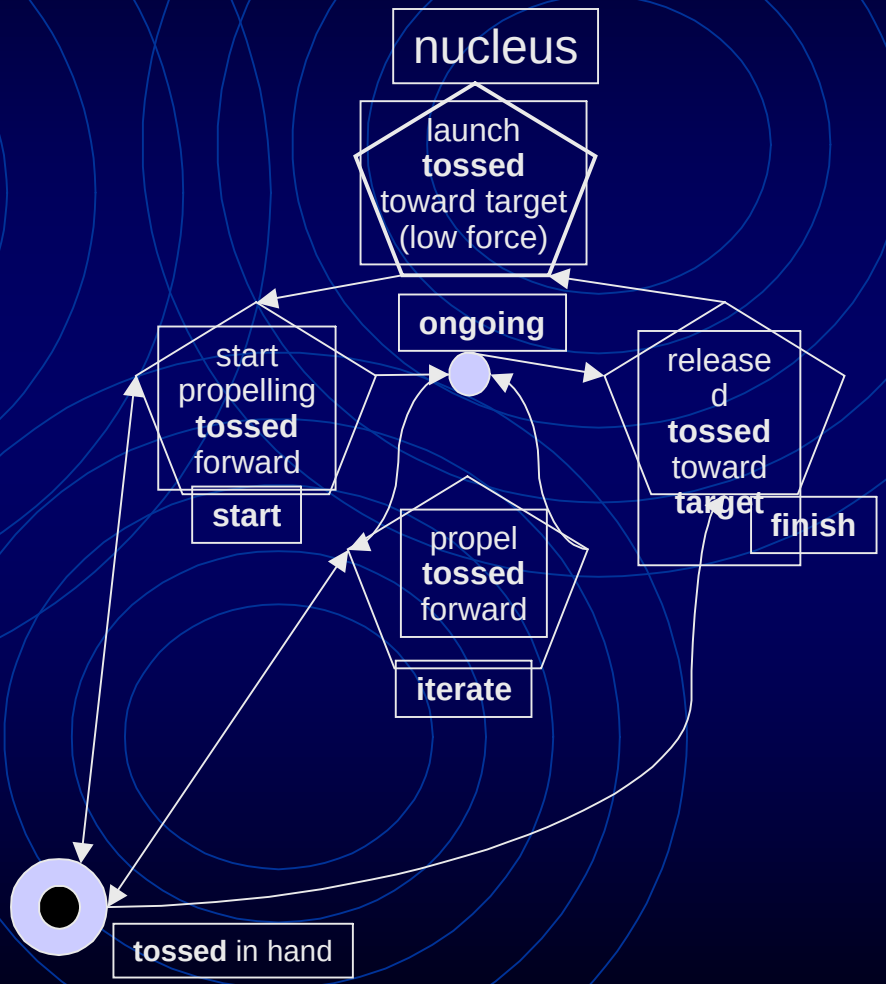
Enabling arcs indicate a static relationship which a transition requires but does **not consume tokens**

Tossing: percept vector



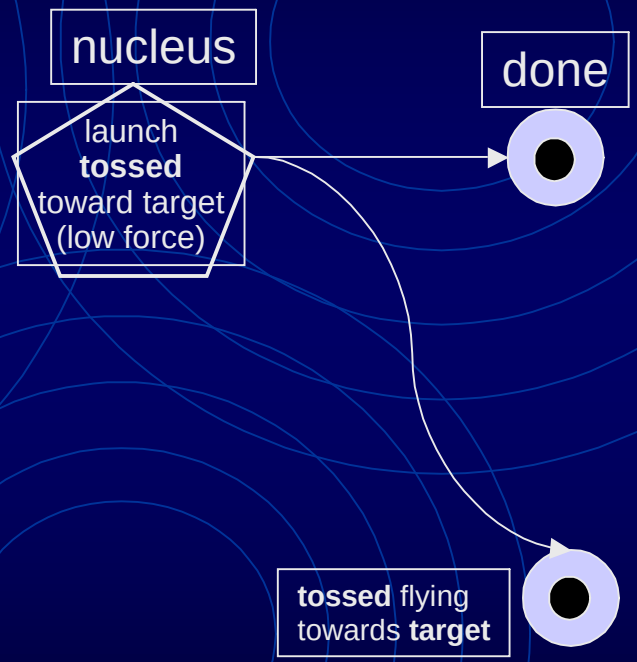
x-schemas

- The firing of this transition causes the execution of the subordinate sequence of actions



x-schemas

- Once these have completed, the transitions firing is complete and tokens are replaced in its output places, asserting that the tossed object is now on its trajectory.
  - The overall token movement can be interpreted as the expenditure of energy in a movement that results in the tossed object leaving the tosser's hand and flying through the air



x-schemas: conclusion

- Have just the right properties needed to drive simulations in the framework
  - X-schemas **can capture fine grained features** of **complex events in dynamic environments**
  - They can be **parameterized** according to different event participants
  - **Constructions** can thus **access the detailed dynamic knowledge** that characterized rich embodied structures merely by specifying a limited set of parameters.
  - Tight **coupling of perception and action** allows highly context sensitive interactions, with the same x-schema producing strikingly different executions based on only slight changes in the percept vector (or in the specified parameters)

Simulation-based inferences

TRANS.ready		SPEAKER does not have a drink
TRANS.nucleus		MARY exerts force via TOSS
	TOSS.enabled	DRINK in reach of MARY
	TOSS.ready	DRINK in hand of MARY
	TOSS.necleus	MARY launches DRINK toward SPEAKER
		MARY expends energy (force-amount = low)
	TOSS.done	DRINK flying toward SPEAKER
		DRINK not in hand of MARY
TRANS.nucleus		MARY causes SPEAKER to receive DRINK
TRANS.done		SPEAKER has received DRINK

back

Empirical part

Data Analysis

"When the going gets tough, the tough get empirical."

Jon Carroll

back

Data analysis

Corpus: British component of the ICE collection
- small (1-2 milllion words), but richly annotated,
(soon even equipped with corresponding sound files)

Data: (finite and non-finite) relative clause constructions
N ~ 1000 across modalities, registers

Procedure: Code manually the set of constructions with respect to grammatical and conceptual features (**parameters**) discussed in the literature

back

Parameters

Type of RCC:

He also said that the Human Rights Act, [which comes into force on October 2], will mean that the Official Secrets Act will have to comply with Article 10.
(**appositive**)

[What I like best] is football
(**nominal**)

The lark is on the wing, [which is normal for larks].
(sentential relative clause)

Parameters

Finiteness of RC: +/-

The man [who mistook his wife for a hat] ...
Things [to do in denver] when you are dead.

Parameters

Type of relativizer:

that, pronoun {a,b,c,...n}, zero

Parameter

Adverbial Sem. Type	+/- classification??
Animacy of OBJ	Silverstein

Parameters

Definiteness of head: indef [art, general term]

def

[art, DEMprn]

Animacy of head: animacy hierarchy

(Silverstein 1976)

Superlative head: +/-

Length of head (in words)

Parameters

Syntactic role of R:	A, S, O, Obl,...
Syntactic role of head:	A, S, O, Obl,...

Parameters

Copular main clause	+/-
Length of ObjRCs VP	(in words)
Subject of ObjRCs	Silverstein (1976)

So, I have not considered intonational parameters.

Some areas of investigation

I. Investigation of recognized RC types
(e.g. S-relatives, O-relatives, etc...)

Question: How do they differ when characterized be means of the set of features mentioned before?

Some areas of investigation

II. **Center embedded** vs. Right-branching structures

Peter who likes books hates John

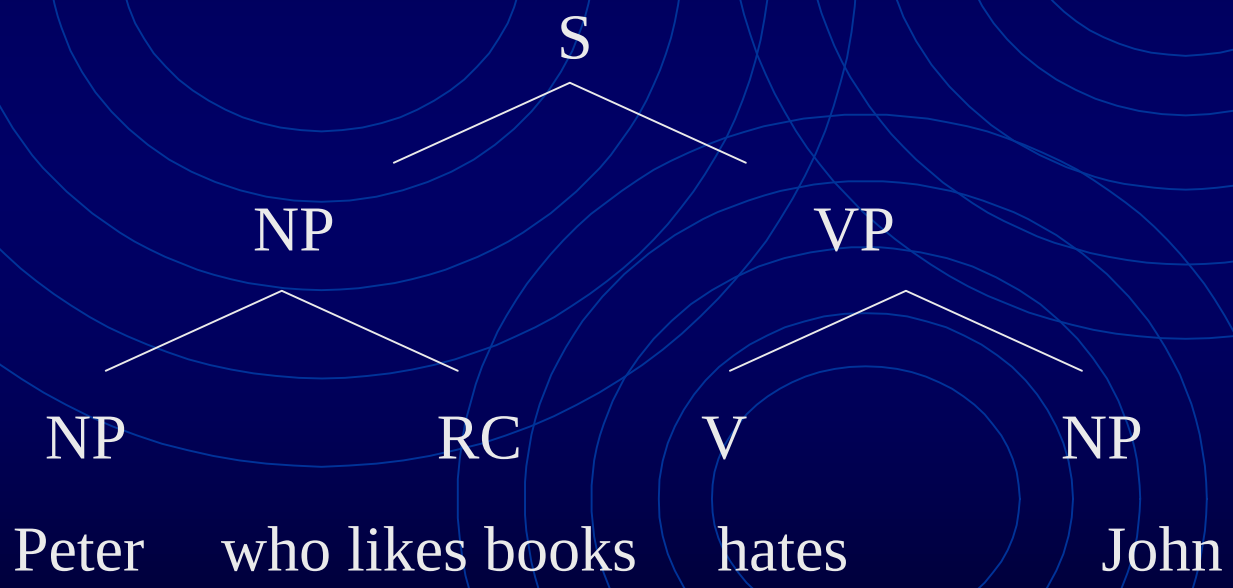
Some areas of investigation

II. Center embedded vs. **Right-branching structures**

Peter hates John who likes books

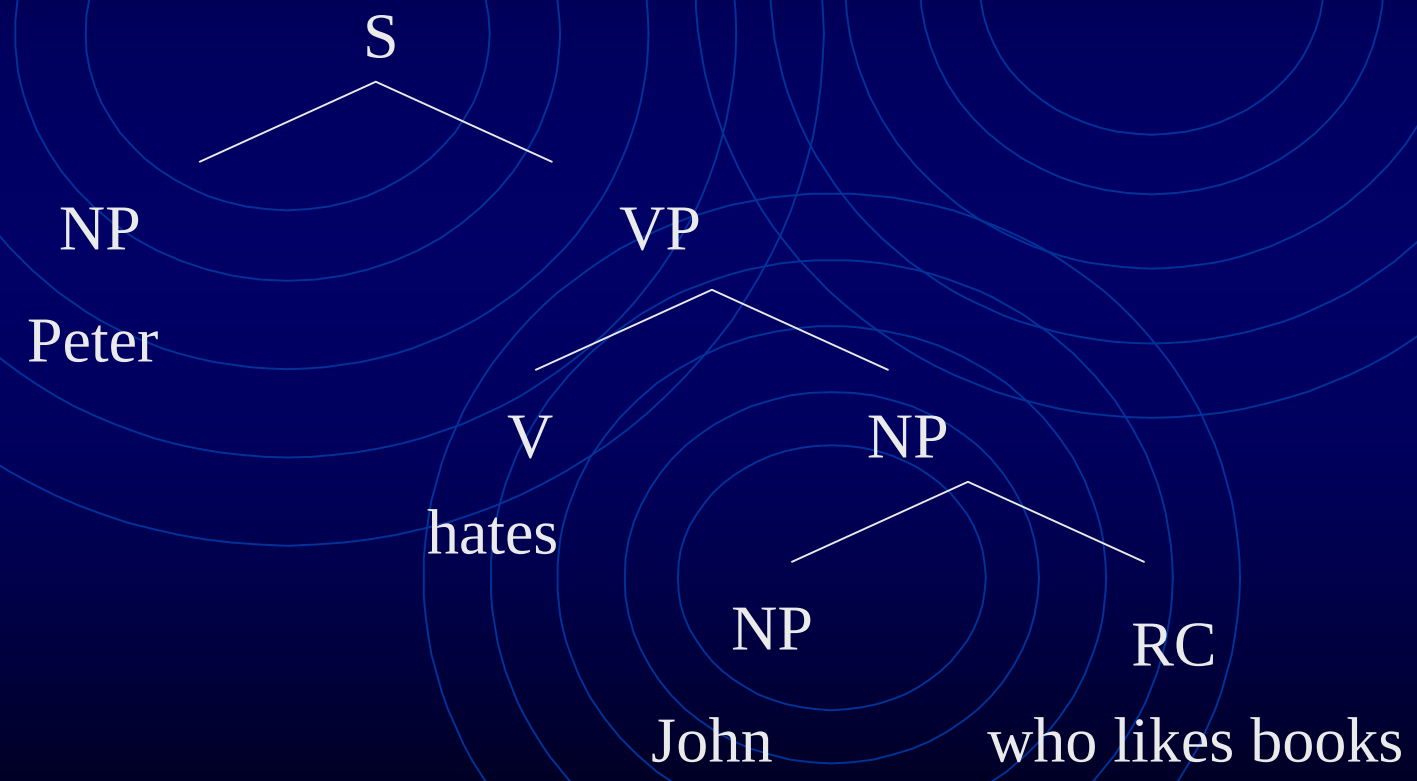
Some areas of investigation

II. **Center embedded** vs. Right-branching structures



Some areas of investigation

II. Center embedded vs. **Right-branching structures**



Some areas of investigation

II. Center embedded vs. **Right-branching structures**

Are there (sets of) RC types that occur predominantly in one of the two structures and if so, why?

Effects on processing?
Usually, it is claimed that O-relatives are harder to process than S-relatives, but maybe this may not be true as general statement; maybe some sub-type cause less trouble?

Some areas of investigation

III. Comparison of similar RC types

i. This is the question, **to** which John does not know the answer. (**pied piping**)

ii. This is the question, which John does not know the answer **to**. (**dangling preposition**)

iii. This ever-changing world in which we live in (**preposition copying**)

Question: Under what conditions are speakers more likely to produce one variant than the other?

Note: This question alone could be the topic of a dissertation

(cf. Gries 2000 on particle placement)

Data analysis

How should the data be evaluated?

Analysis: Given the system needs to make a nominal choice...

-monofactorial analysis

-assess absolute strenth of variables in terms of correlation coefficients

-pairwise comparison of factors

-identify relative strength of variables

-multivariate techniques

-psychologically more plausible: all variables enter decision tree (but how?)

-General Linear Model, e.g. **discriminant analysis**?

-CART (**C**lassification and **R**egression **T**rees) **analysis**?

Data analysis

How should the data be evaluated?

Analysis: Given we want to find some objective measure dealing with the enumeration of constructions problem...

- Hierarchical configurational frequency analysis?

- Cluster analysis?

- Analogical modeling? [tiMBL: Tilburg machine based learning]

back

A parsing hypothesis

In languages like English, in which the external head of the construction precedes the clausal pattern, we might be able to predict much of the structural pattern to come (i.e. type of RCC) on the basis of linguistic and/or conceptual properties of the head

So, the statistical analysis will be designed such as to observe correlations between head features and RC features.

Aside:
If this hypothesis is correct, one would expect to find differences across languages that share the head initial parameter, but which differ with respect to how they conceive of the referent of the head NP.

back

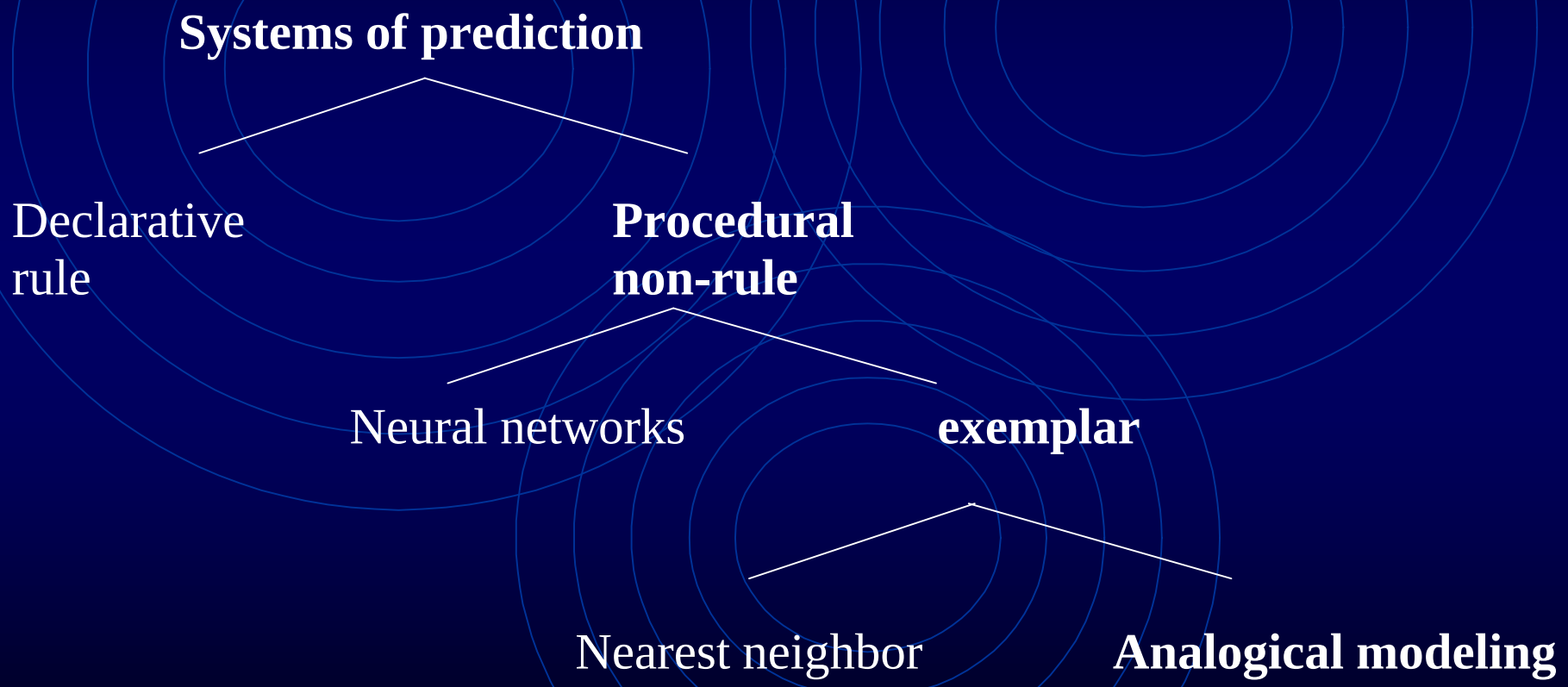
goto AM

A little something on Analogical Modeling

" …the faculty for perceiving analogies is the best indication for genius." William James

back to emp back

Different types of language prediction



given contexts, subsets, suprasets

- Predictions are based on dataset of occurence
- each occurrence is specified in terms of a set of variables
- Start: a **given context (GC)**
  - **(**variables are sprecified, no outcome is given)
- Task: predict outcome for given context in terms of occurrences in dataset
  - cf. example 'spelling to pronunciation'
- 0 < m (number of specified variables) < n (number of variables)
  - In STP example m=3

Example

predicting pronunciation from spelling simplified data set

outcome	variables	specification
k-c	ake	cake
k-c	all	call
k-c	an0	can
k-c	ar0	car
k-c	at0	cat
s-c	ell	cell
s-c	ent	cent
s-c	ert	certain
tsch-c	hec	check
k-c	los	close
k-c	lou	cloud
tsch-c	hin	chin
s-c	irc	circle
s-c	irc	circus
s-c	ity	city
k-c	lam	clam
k-c	los	close
k-c	oat	coat
…	…	…
s-c	ymb	cymbal

given contexts, subsets, suprasets

- For each subset of of variables defined by the given context, we determine which occurrences in the dataset occur with that **subset**
  - **cf. data set example**
- Each subset of variables is called a **supracontext (supraC)**
- We have a total of 2^m **supracontexts**
  - **e.g. STP example given context *cake*: m=3 --> 8 supracontexts**
    - **ake** — **supraC of cake**
    - **ak-** — **all three**
    - **a--** — **two variables, one ignored**
    - **-ke** — **one variable, two ignored**
    - **--e**
    - **-k-**
    - **a-e**
    - **---**

SupraC: deterministic vs non-deterministic

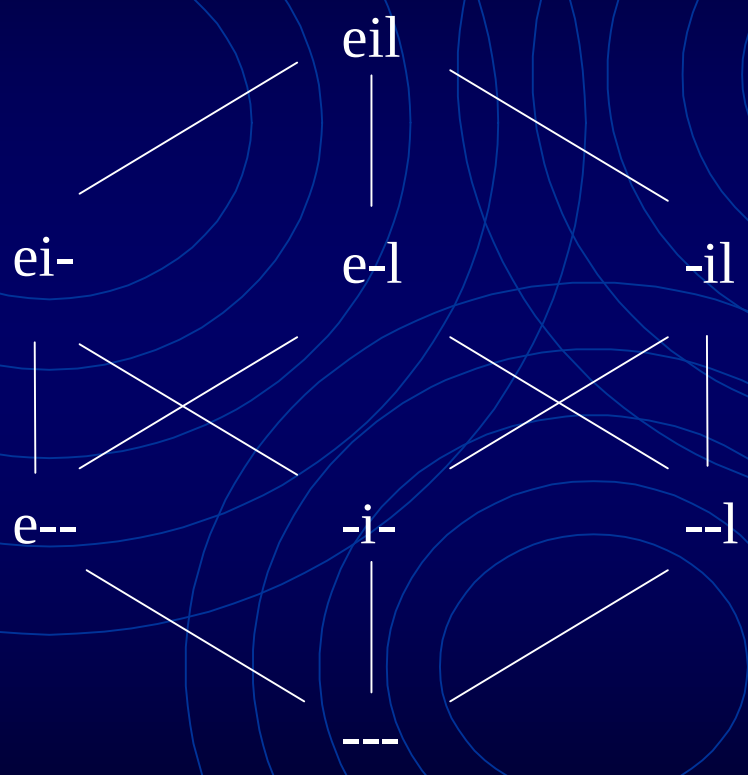
- There are only two types of homogeneous supraC:
  - Deterministic supraC
  - Non-deterministic supraC

For each supraC, frequencies of occurence in the datasets are collected

•	k-c	s-c	tsch-c	
• eil	-	-	-	
• ei-	-	-	-	
• e-l cell	-	1	-	homo, det
• -il	-	-	-	homo, det
• e-- cell, cent, certain	-	3	-	homo, non-det
• -i- chin, coin	1	-	1	
• --l call, cell, cycle, cyclone	1	3	-	
• --- <whole dataset>	21	9	3	

Partially ordered lattice

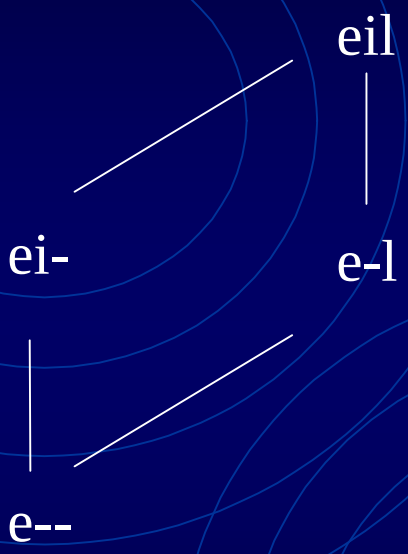
Example:
ceiling



By following lattice upwards, we can determine for each supraC the set of subC

SubCs of supraC e--:

Example:
ceiling



By following lattice upwards, we can determine for each supraC the set of subC

Homogeneity (and heterogeneity)

- A supraG is **homogeneous if all its possible subcontexts behave identically**
- In predicing the outcome for a GC, we **only apply information found in homogeneous supraC**
- All Heterogeneous supraC are ignored
- Problem: **determine homogeneity of each supraC** defined by a given context

Homogeneity (and heterogeneity)

- We determine whether a supraC is homogeneousby using a **nonlinear statistical procedure based on measuring the number of disagreements between different occurences within a supraC**
- **If no subcontext of the supraC increases the number of disagreements, the supraC is homogeneous**
- --> end up minimizing the number of disagreements in a supraC

deterministic vs non-deterministic homogeneous supraC

- Deterministic: only one outcome occurs
- non-deterministic: all occurrences occur with only one subC of the supraC
- Example *ceiling : 2 deterministic supraC*
  - *e-l and e--*
- ***The more general supraC is homogeneous because it contains only examples of the s-c outcome***
  - ***e-- is deterministic (no subcontext od e-- behaves different than e--)***

back