# On Eliciting Requirements For Agent Semantics From Linguistics Concepts

E. M. Sherwood

Distinguished Member of the Technical Staff

*Motorola Labs*
*602-952-4001*

*Everett.M.Sherwood@Motorola.Com*

*Chair, Semantics Technical Committee, FIPA*

# Semantics is a limiting issue for agent transactions.

- We use agreed protocols for the exchange of data.  On each side of the exchange, data are held within a model of a domain.

- BUT,
Agents lack agreement on
_the methods for semantic exchange._

- BECAUSE
Semantic exchange is not defined

# An Analogy for a Goal …

# Agent-to-Agent ESPERANTO

# Agent-to-<u>Human</u> ESPERANTO

# We are Seeking
## An Engineering Solution

# We Will Build It.

# An Engineering Solution is Formed From …

- Precise <u>definitions</u>
  - Including: Words, sentences, ontologies, meaning, semantics, knowledge, …
- Underlying <u>principles</u>
  - Equality, replacement,
- Embodiment
  - Key data <u>structures</u>
  - Needed <u>algorithms</u>
- The Criteria for Success  and
  A Means to Establish Proof of Success

# Perspective

- Agent Semantic Transactions is <u>Not a solved problem</u>
- As a first step towards reaching a solution, a clear <u>statement of requirements</u> is needed.
- Because natural language is our primary source of semantic exchange, <u>linguistics offers a source of those requirements</u>
- Within linguistics, semantic exchange may be partitioned into two levels
  - Low-level; referents, structures, sentences
  - High-level; dialogues, Q&A, roles/identities, speech acts, …
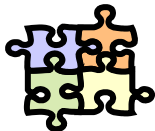- High level semantics depends on low level semantics

# Other Sources of Requirements

- Sign Language
- Mathematics
- Logic
- Maps
- Signs / Semiotics
- Chemistry
- Music
- Body Language

# To instigate discussion, included in the presentation are:

- Some <u>issues</u> in semantic transactions
- Some <u>draft requirements</u>
- Some <u>claims</u> and some <u>local</u> <u>definitions</u>

  - Slides marked with  <u>reflect my personal perspective</u>

# Some cautions for using linguistics as a source of requirements

- Only a small part of linguistics, (semantics and grammar) offers support
    - Phonology, History, Semiotics, … provide little additional information
- There are substantial disagreements among linguists on many topics
    - "The Linguistic Wars"
- Much of the work is speculative and has little experimental validation

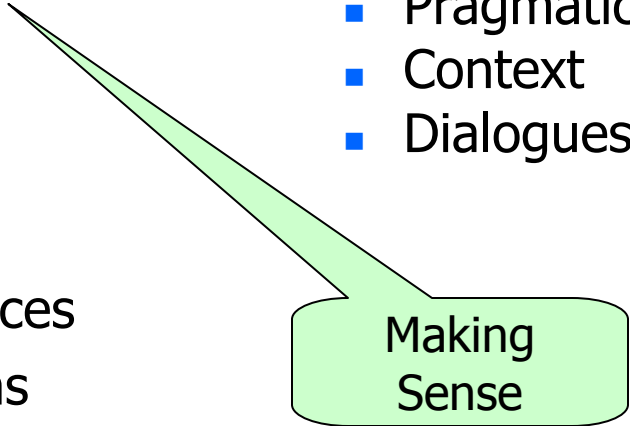- *However, linguistics references supply <u>excellent sets of posed problems, insights, and examples</u>*
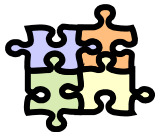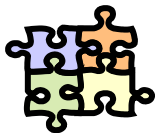
# Coarse Outline

- Why Agents Need Semantics
- Requirements
- A Perspective on Semantics

- Low Level Semantics
  - Referents and Words
  - Lexical Structures
  - Sentences
  - Organization of Sentences
  - Control of Organizations

Making Sense with Others

- High Level Semantics
  - Questions & Answers
  - Identities and Roles
  - Pragmatics
  - Context
  - Dialogues

Making Sense

A Definition to Start ….
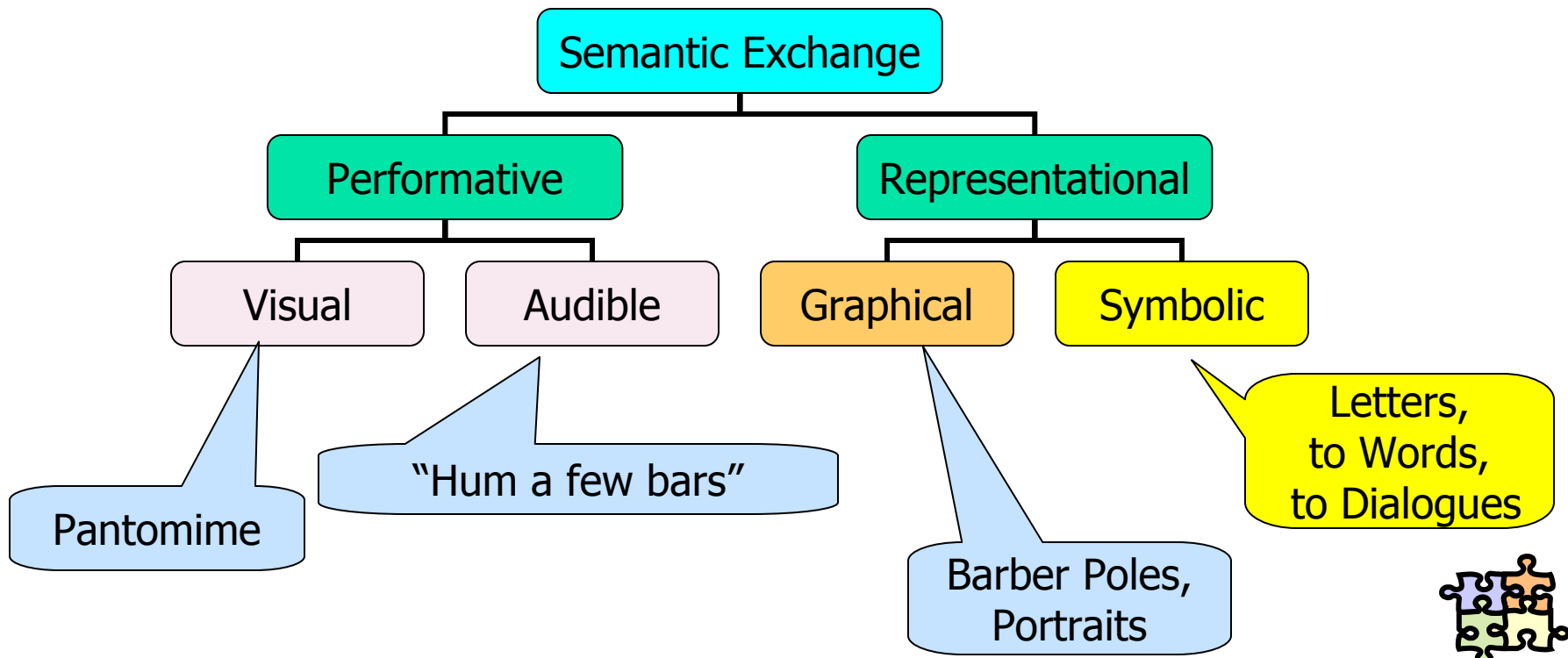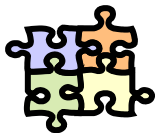
# Semantics

An account of
how abstracted representations
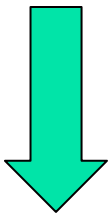are linked to
their real world counterparts.

*-- After Patrick Hayes, 1974*

# How We Inform Each Other

```
                    ┌─────────────────────┐
                    │  Semantic Exchange  │
                    └─────────────────────┘
                    ┌──────────┴──────────┐
          ┌──────────────┐        ┌──────────────────┐
          │ Performative │        │ Representational │
          └──────────────┘        └──────────────────┘
          ┌──────┴──────┐         ┌──────────┴─────────┐
     ┌────────┐  ┌─────────┐  ┌───────────┐    ┌──────────┐
     │ Visual │  │ Audible │  │ Graphical │    │ Symbolic │
     └────────┘  └─────────┘  └───────────┘    └──────────┘
```

Pantomime

"Hum a few bars"

Barber Poles, Portraits

Letters, to Words, to Dialogues

# Linguistic Viewpoints on Meaning

Mostly Words

- Usage of words (Wittgenstein)
- Referents
- Field (related terms) (Muller)
- Affect (Osgood)
- Prototype / Ideal / Category (Lakoff)
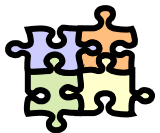- Semantic nets (e.g., Schank)

Mostly Sentences

- Entailments  (Logic)
- Translation
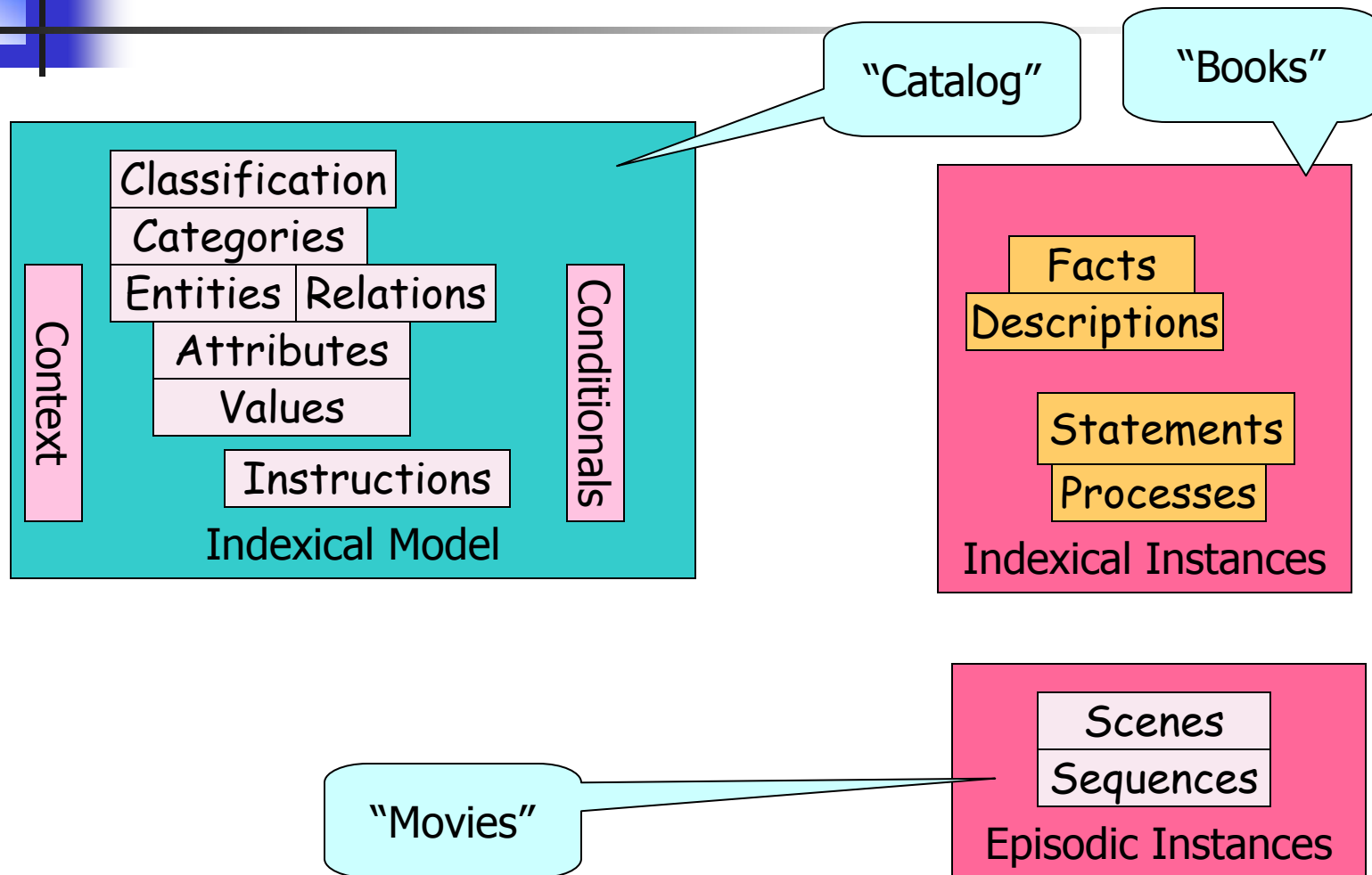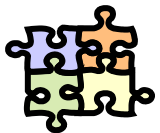- Process / Action (meaning exists in terms of process)

# Semantics is not knowledge

- Knowledge (what-we-know)
  is captured <u>in terms of</u> semantics
  and held as <u>instances</u> in <u>memory.</u>

- Better representation and expressivity in semantics
  enable better <u>recording</u> of knowledge.

- <u>Domain models</u> mimic memory,
  the recorded form of knowledge
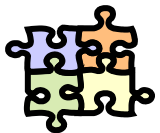  - Lexical
  - Indexical
  - Episodic

# Knowledge in Domain Models (Memory Structures)

"Catalog"

"Books"

**Indexical Model**
- Classification
- Categories
- Entities  Relations
- Attributes
- Values
- Instructions
- Context
- Conditionals

**Indexical Instances**
- Facts
- Descriptions
- Statements
- Processes

"Movies"

**Episodic Instances**
- Scenes
- Sequences

# Two fundamental Low-Level Semantics Requirements

- Representation
  - Can we say what we mean?

- Expressivity
  - How well can we say what we mean?

# Low-Level Semantics IS about representation and about expressivity BUT it is not about "TRUTH."

Knowledge may be …

- Wrong
- Redundant
- Conflicting
- Missing
- Degraded

- Duplicated
- Changing
  - Changed again
- Validated (or NOT)
- Referenced (or NOT)

In 786 AD, it was "true" that the world was flat.

BEFORE we can assess "TRUTH" of the matter, we need to express it.
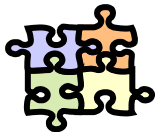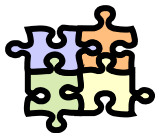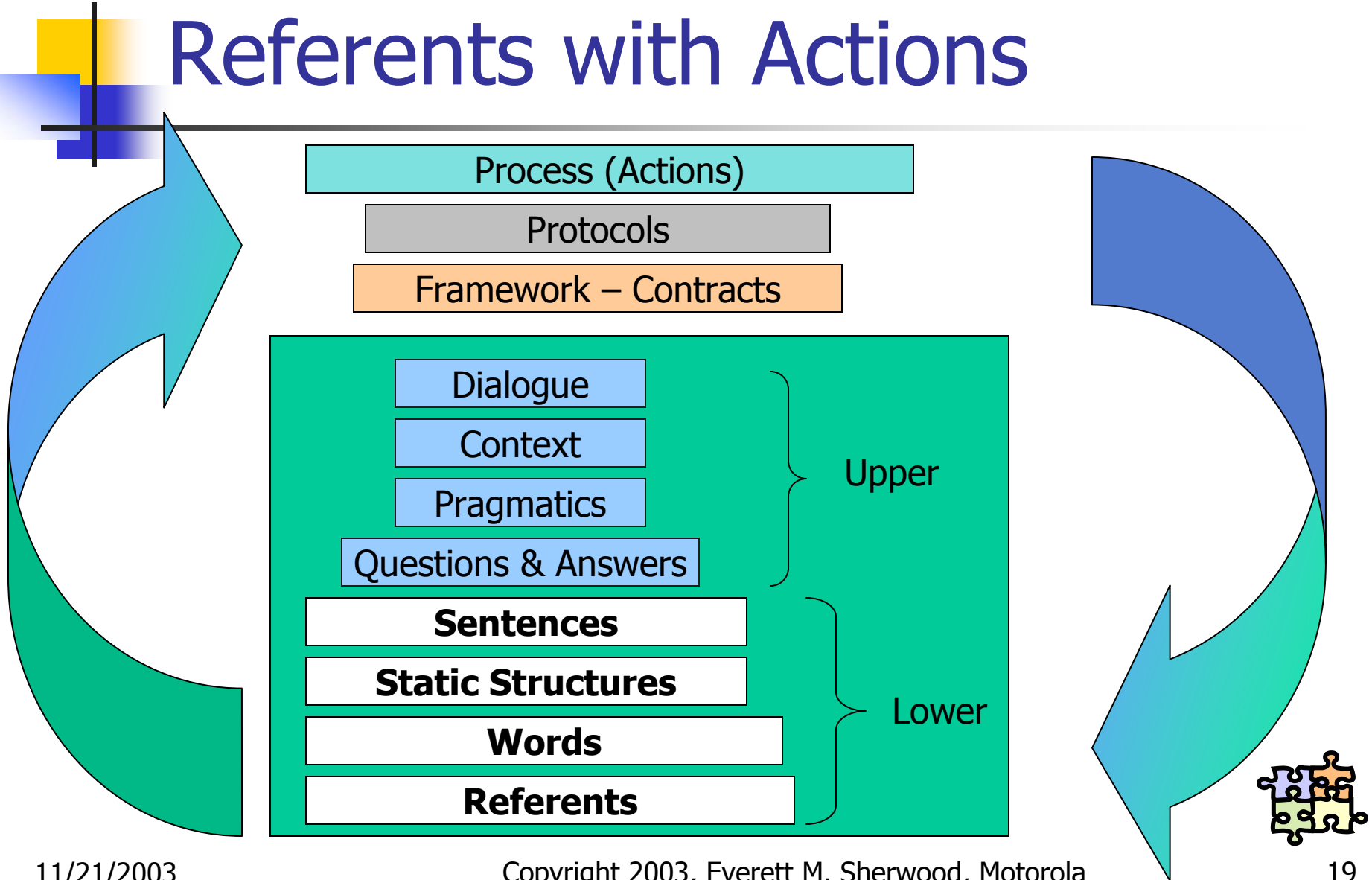
# The Goal is Not to Reason

- True/false are not of interest…

- The goal is to represent so that some <u>external</u> process may reason.

If you want modal logic, we must provide modal verbs;
If you want temporal logic, we must provide tense and aspect.
If you want subtraction, we must provide numbers.

# Meaning Links
# Referents with Actions



Process (Actions)

Protocols

Framework – Contracts

Dialogue

Context

Pragmatics

Questions & Answers

**Sentences**

**Static Structures**

**Words**
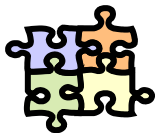
**Referents**

Upper

Lower

# Semantic Sense vs. Meaning

- George: "Please give me the hammer."
- Gracie: "There is ice in the cooler."

**Makes sense
but is not meaningful to the process**

**Makes no sense**

Colorless green ideas sleep furiously.
-- Noam Chomsky
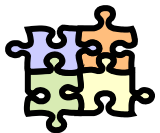
# Low Level Semantics Makes Sense

- *Low-level Semantics is about making sense* by placing the right words into sentence structures.

- Low-level Semantics
  - Is Not about Reasoning
  - Is Not about Truth
  - Is Not about Knowledge
  - Is Not about Meaning

# Meaning is based on connections:

I. From real-world <u>referents</u> to their lexemes — What the Words Identify

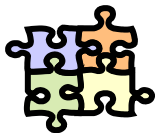II. Through <u>difference relationships</u> of lexemes with other lexemes — The Right Words

III. By the coupling of lexemes into <u>sentences</u> — Making Sense

*Low Level*

IV. <u>To affect processes</u> — Providing Meaning
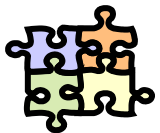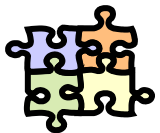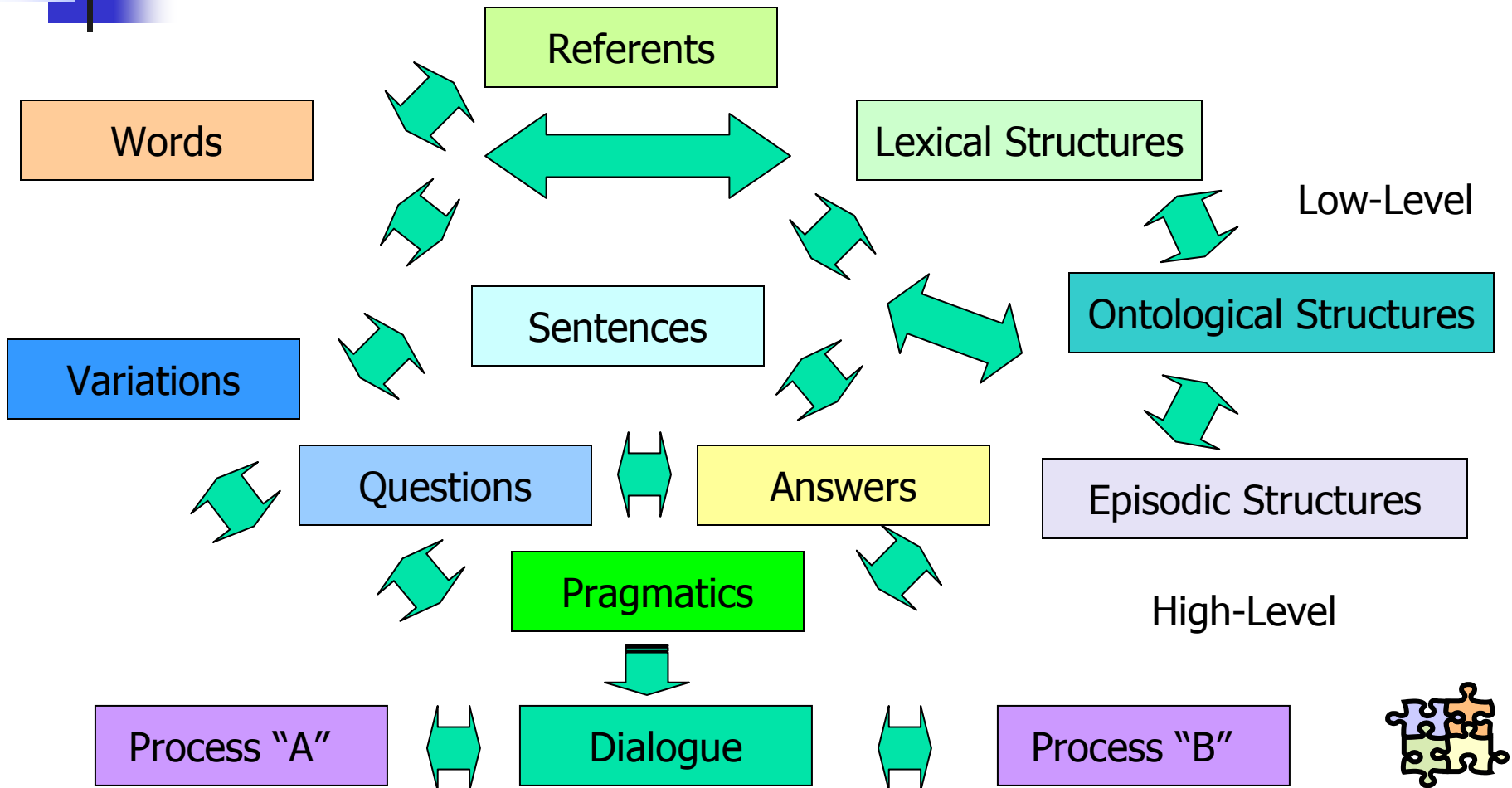
# High-level Semantics Enables Exchange

- Question and Answers
- Context
- Pragmatics
  - Roles & Identities
  - Deixis
  - Implicature
  - Speech Acts
- Dialogue

# The Interactions in Linguistic Semantic Exchange are Complex

Referents

Words

Lexical Structures

Low-Level

Sentences

Ontological Structures

Variations

Questions

Answers

Episodic Structures

Pragmatics

High-Level

Process "A"

Dialogue

Process "B"

Copyright 2003, Everett M. Sherwood, Motorola

# A Difficult Problem

Think of the first time someone considered how to represent integers and real numbers within a fixed word size...
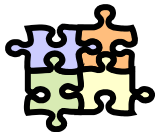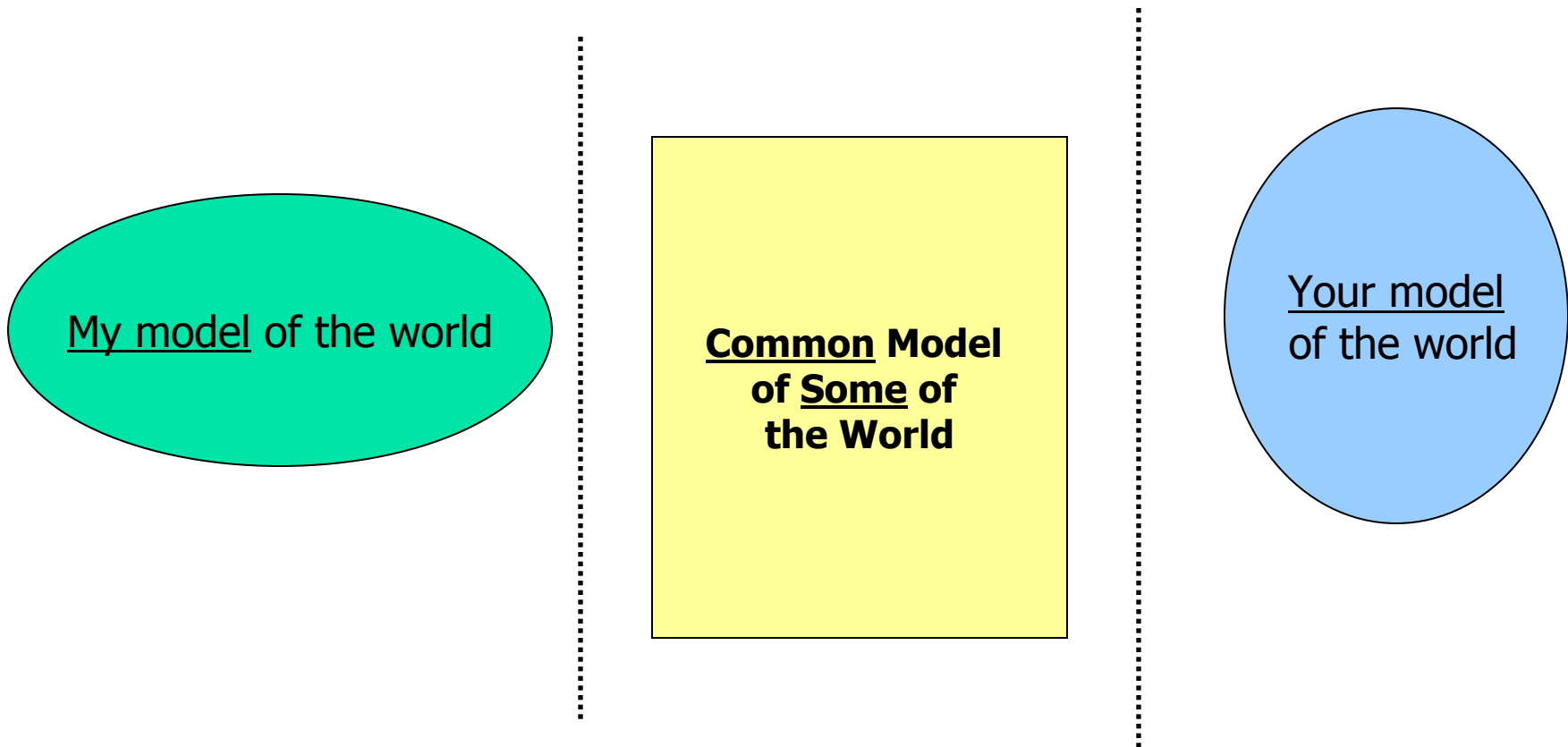
## This is much more complicated.
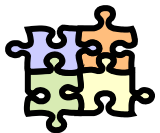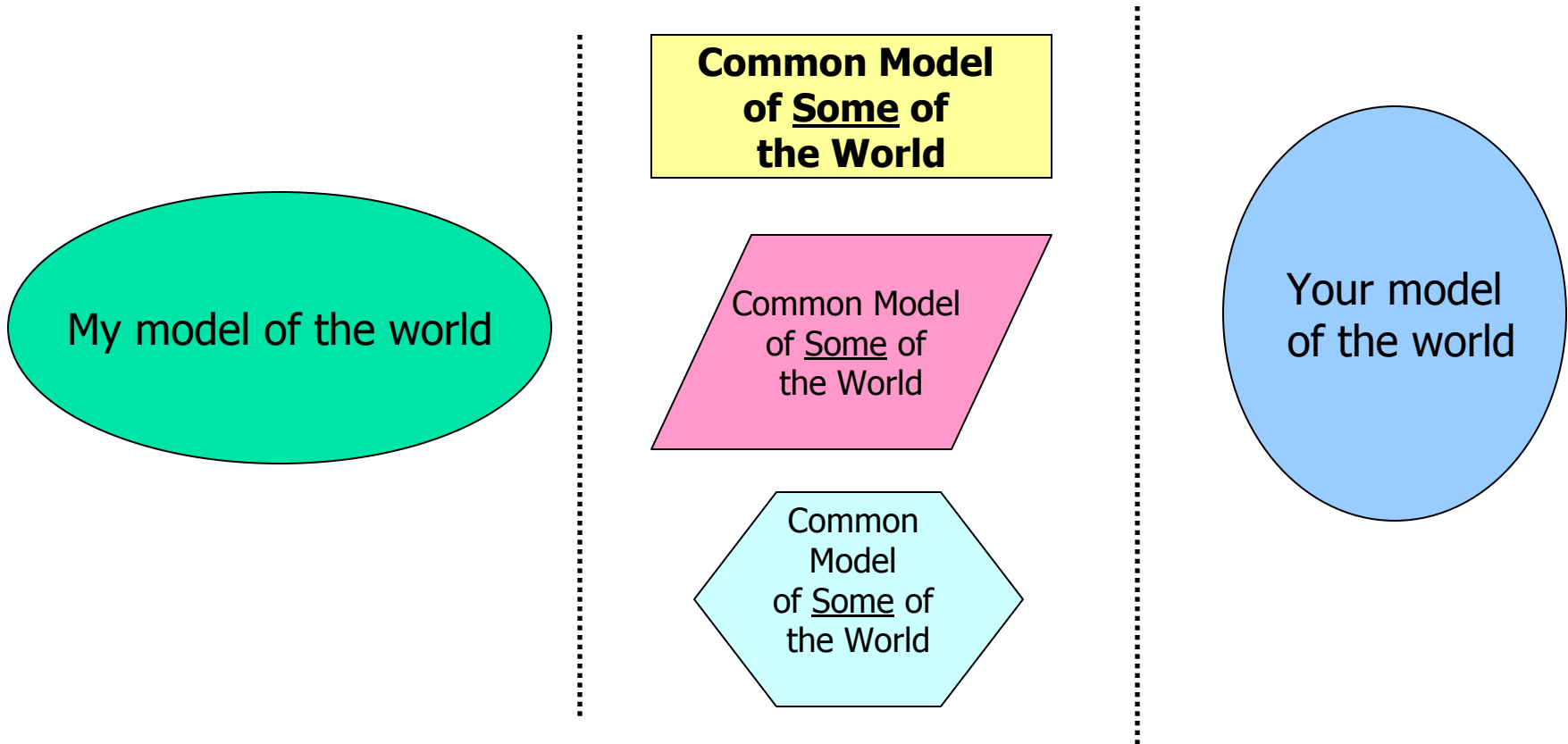
# Part 1 Low-Level Semantics

## Making Sense

# Exchange is enabled and constrained by the common models

My model of the world

**Common Model of Some of the World**

Your model of the world

# There are many common models

**Common Model of <u>Some</u> of the World**

My model of the world

Common Model of <u>Some</u> of the World

Common Model of <u>Some</u> of the World

Your model of the world

# Three ways to acquire knowledge

**Sensors"**

A model of the world

Another model of the world

**Analysis**

**Exchange**
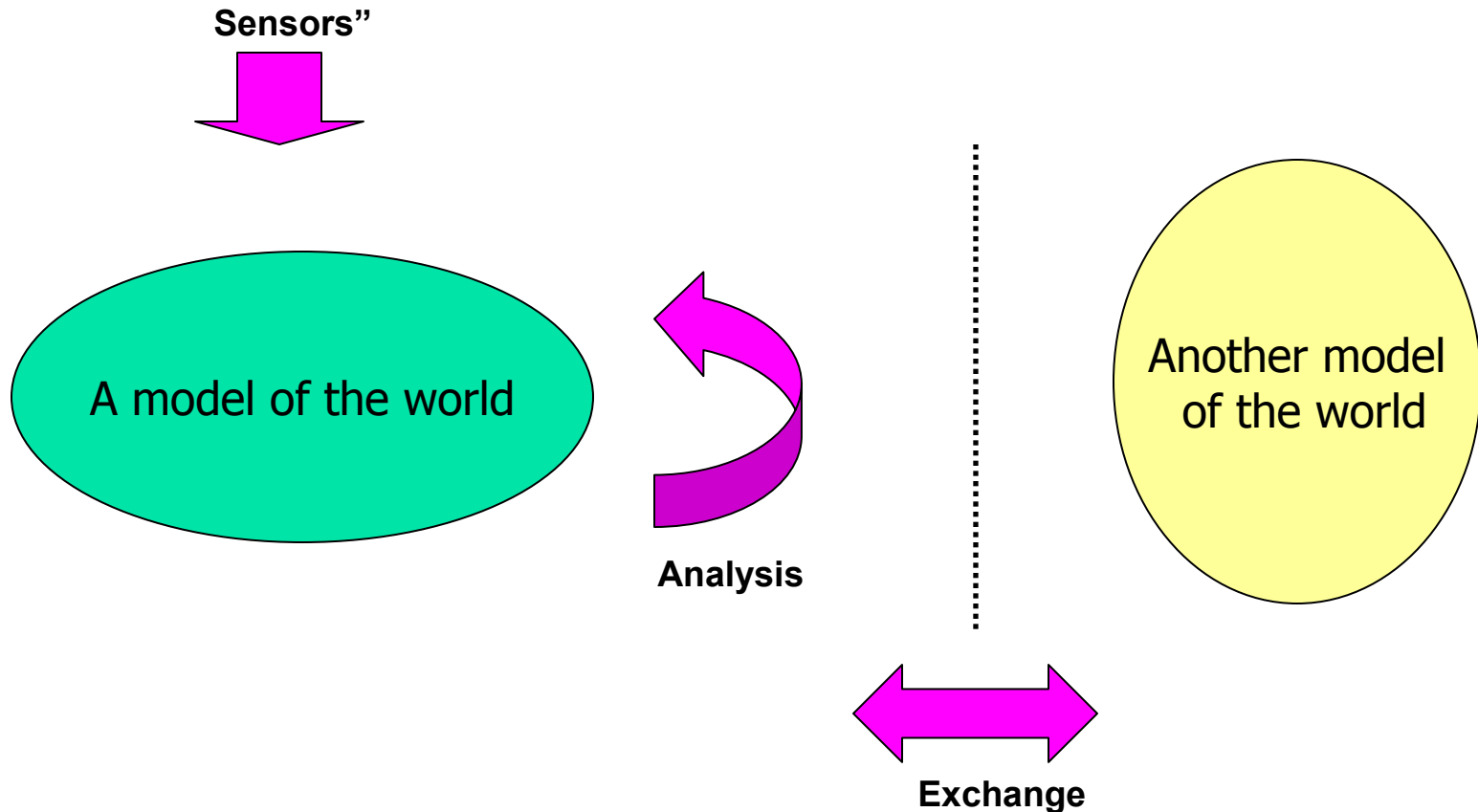
# Inside the Symbolic Model

**Domain Model**

Lexical Structures

Indexical & Episodic Frameworks

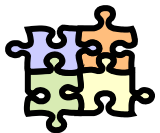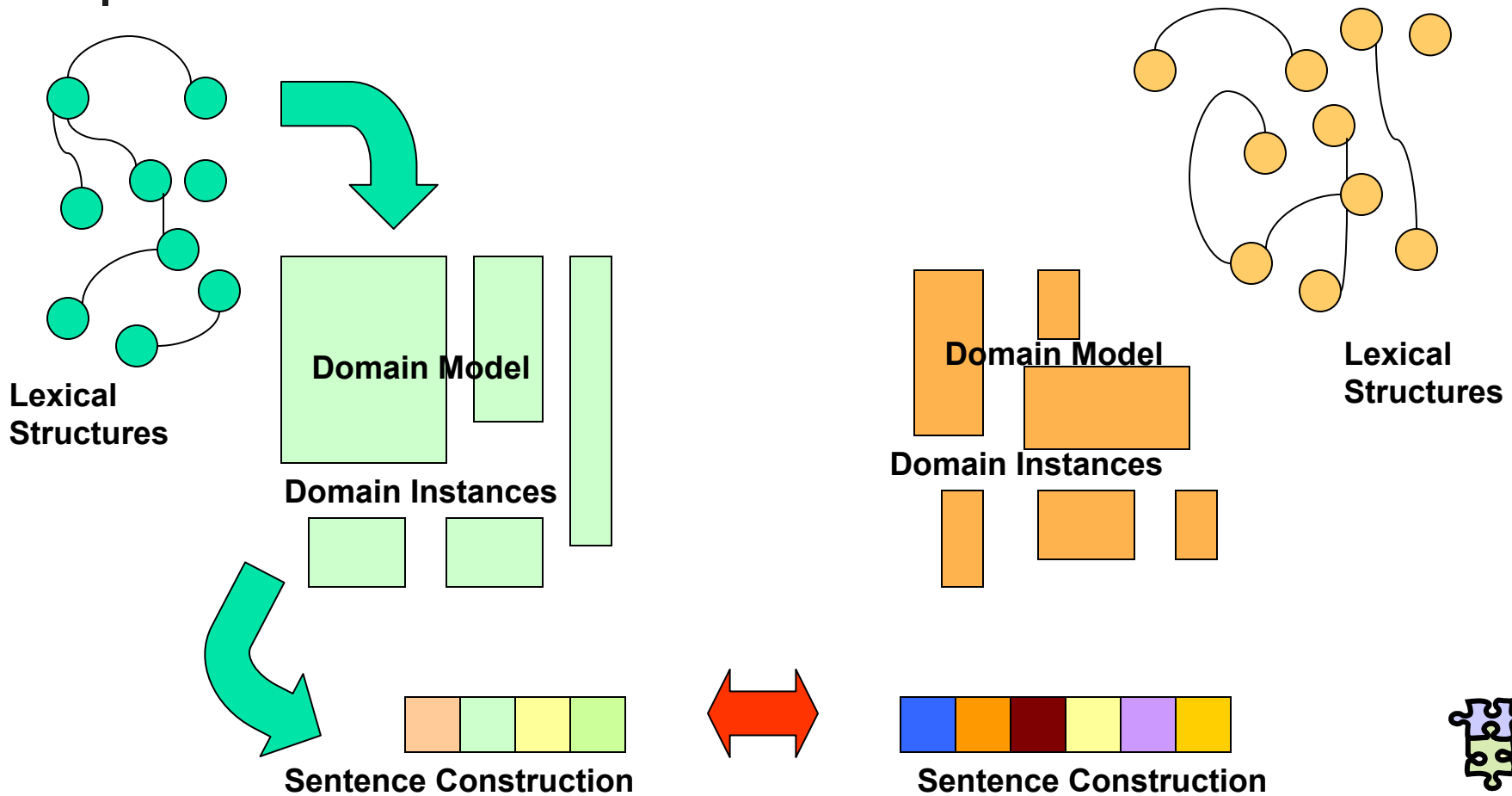Domain Instances
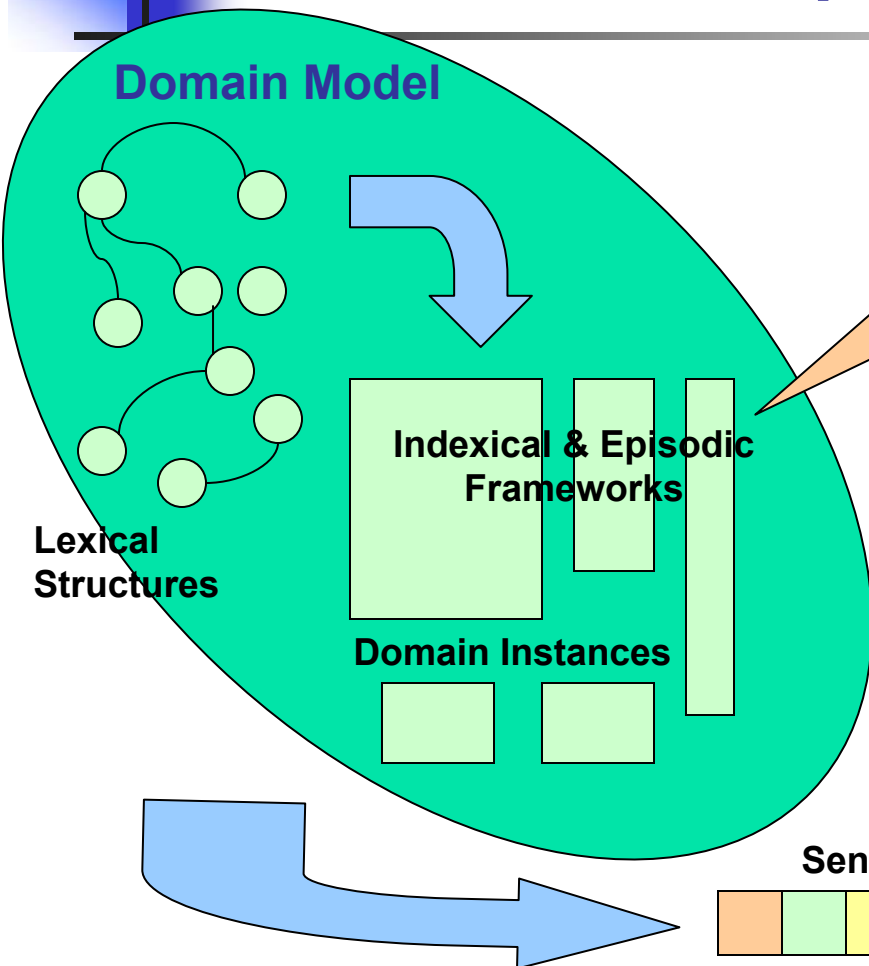
A Personal Model
"What I Know"
Expressed as sentences
and held
episodically and indexically

When we answer a question,
we use the "what we know"
which is held in the domain model to
chose the lexemes needed. Then we
apply a grammar ("how we say it")
to populate the sentences
that comprise the answer.

Sentence Construction

# Exchanges between models are made with sentences



Lexical Structures

Domain Model

Domain Instances

Sentence Construction

Domain Model

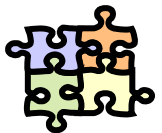Domain Instances

Lexical Structures

Sentence Construction

# Draft Requirements

- Agents will need to map:
  - Lexical Structures
  - Ontological Structures
  - Episodic Structures
- Two possible methods
  - A direct mapping or
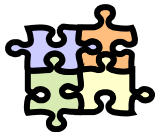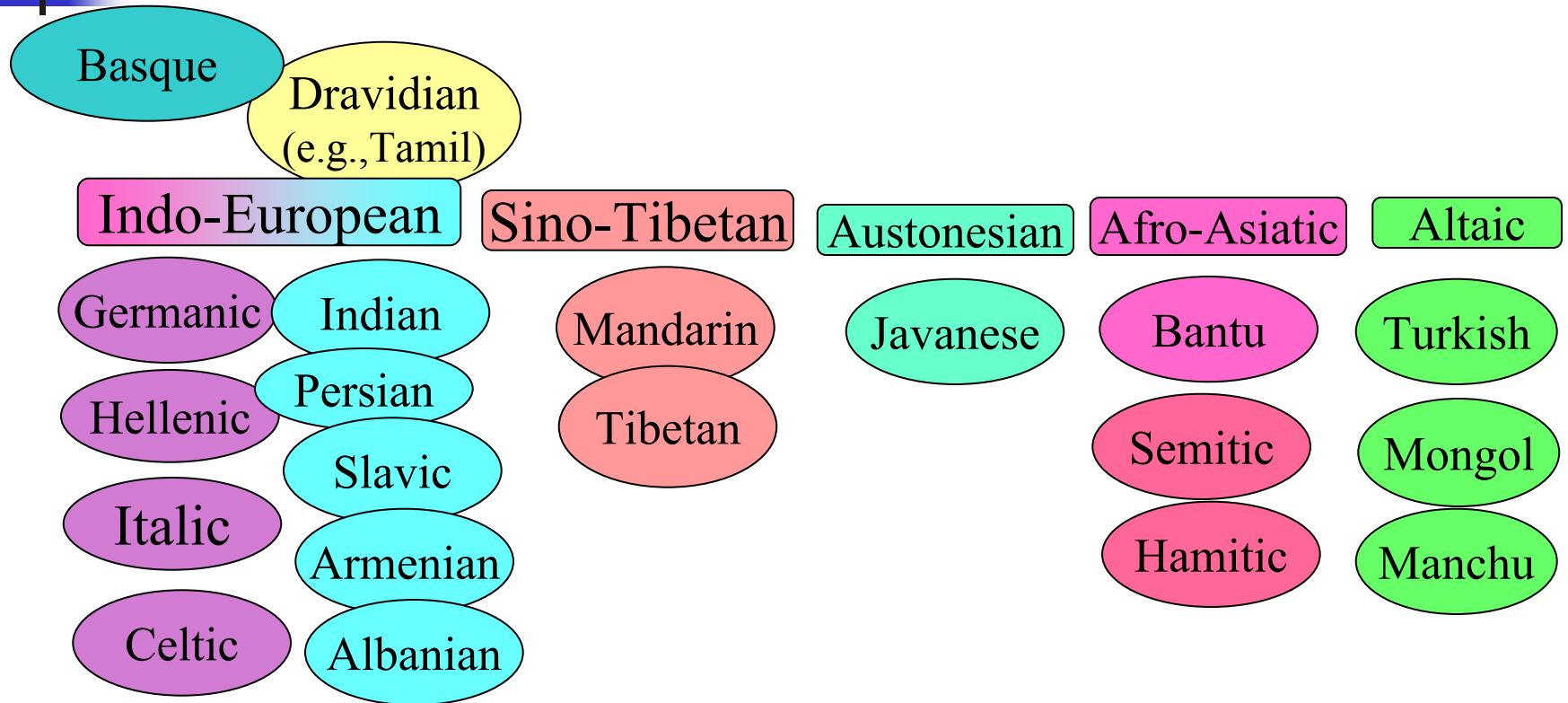  - An indirect mapping to a shared structure

# Linguistics
# as a source of requirements

Natural Language
has been our medium of exchange
for a very, very long time

# Natural Language is a basis of requirements for symbolic semantic exchange.

Basque

Dravidian (e.g.,Tamil)

**Indo-European**

Germanic
Indian
Hellenic
Persian
Italic
Slavic
Armenian
Celtic
Albanian

**Sino-Tibetan**

Mandarin
Tibetan

**Austonesian**

Javanese

**Afro-Asiatic**

Bantu
Semitic
Hamitic

**Altaic**

Turkish
Mongol
Manchu

There are many ways to "say" the same thing …

# Coarse definitions

- A <u>Language</u> consists of <u>a set of words</u> and the set of <u>possible sentences</u> they can create
  - It may also include sounds, gestures

- A <u>Sentence</u> –
  - An output of a grammar
  - What we <u>can</u> say / write

- <u>Lexeme</u> – a element used in a sentence
  - Usually a word or combination (e.g. *blue-green* ocean)
  - A morpheme – an element of a word
    - Un<u>pleasant</u>

# What the Words Mean -- Lexemes and Lexical Structures
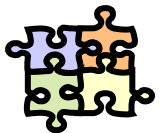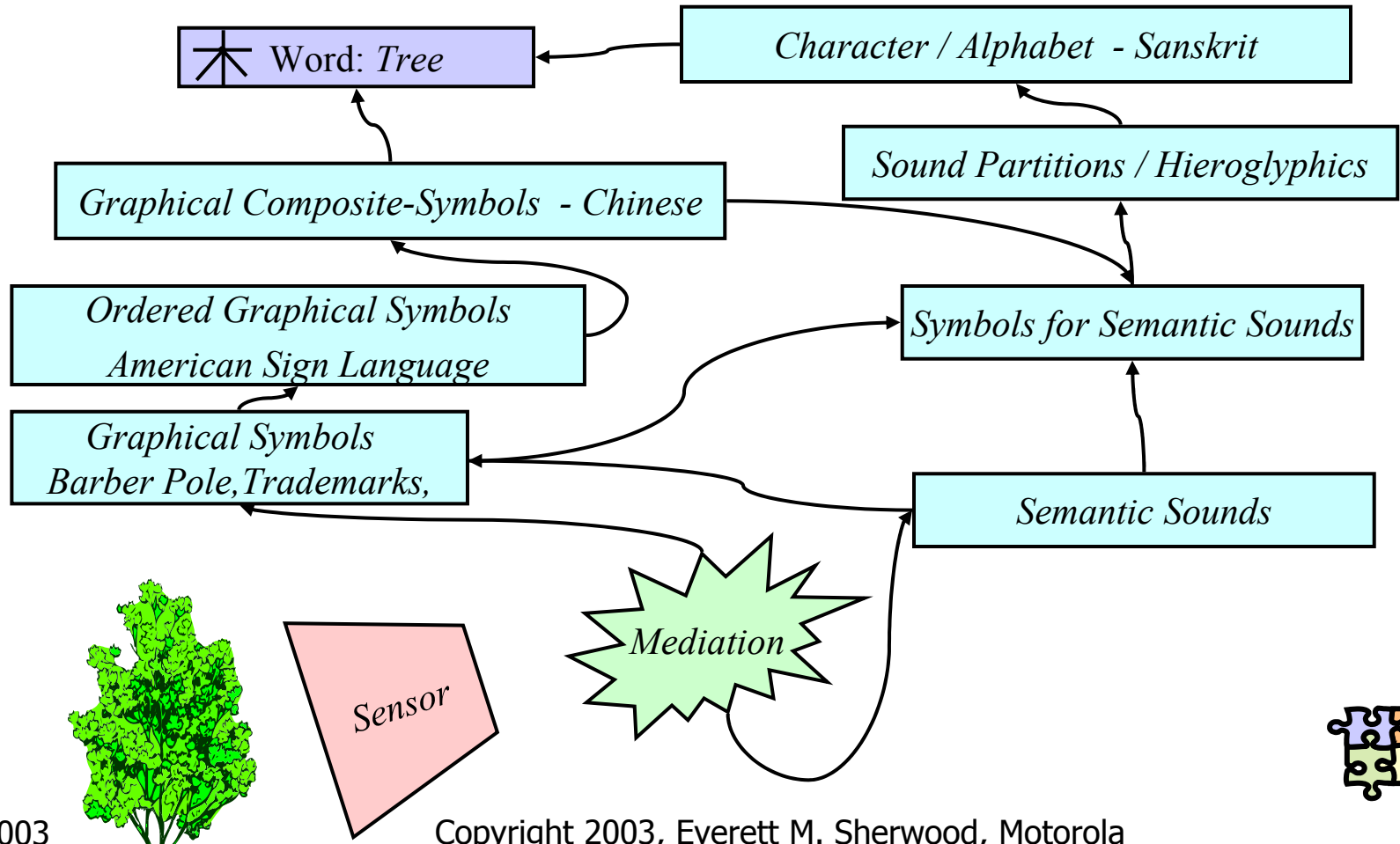
Lexemes

Lexical Structures

# Referents: From Reality to Representation

Word: *Tree*

*Character / Alphabet - Sanskrit*

*Graphical Composite-Symbols - Chinese*

*Sound Partitions / Hieroglyphics*

*Ordered Graphical Symbols American Sign Language*

*Symbols for Semantic Sounds*

*Graphical Symbols Barber Pole,Trademarks,*

*Semantic Sounds*

*Sensor*

*Mediation*

# Referents –
# The real world side of the abstractions

- Referents are:
  - <u>Processes</u>
    - Growing
    - "verb"

  - <u>Results</u> of Processes
    - A Tree
    - "noun"

  - <u>Descriptions</u> that apply to multiple processes and results
    - Slowly (applies to many processes)
    - "Adverb"

    - Green (applies to many results)
    - "Adjective"
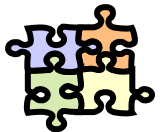
# Referential Semantics

- Every lexeme must be traceable to some real world reference(s).
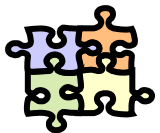
**Blue**:
  the color of the <u>clear sky</u>
  <u>in the day</u>
  viewed <u>from</u> the earth
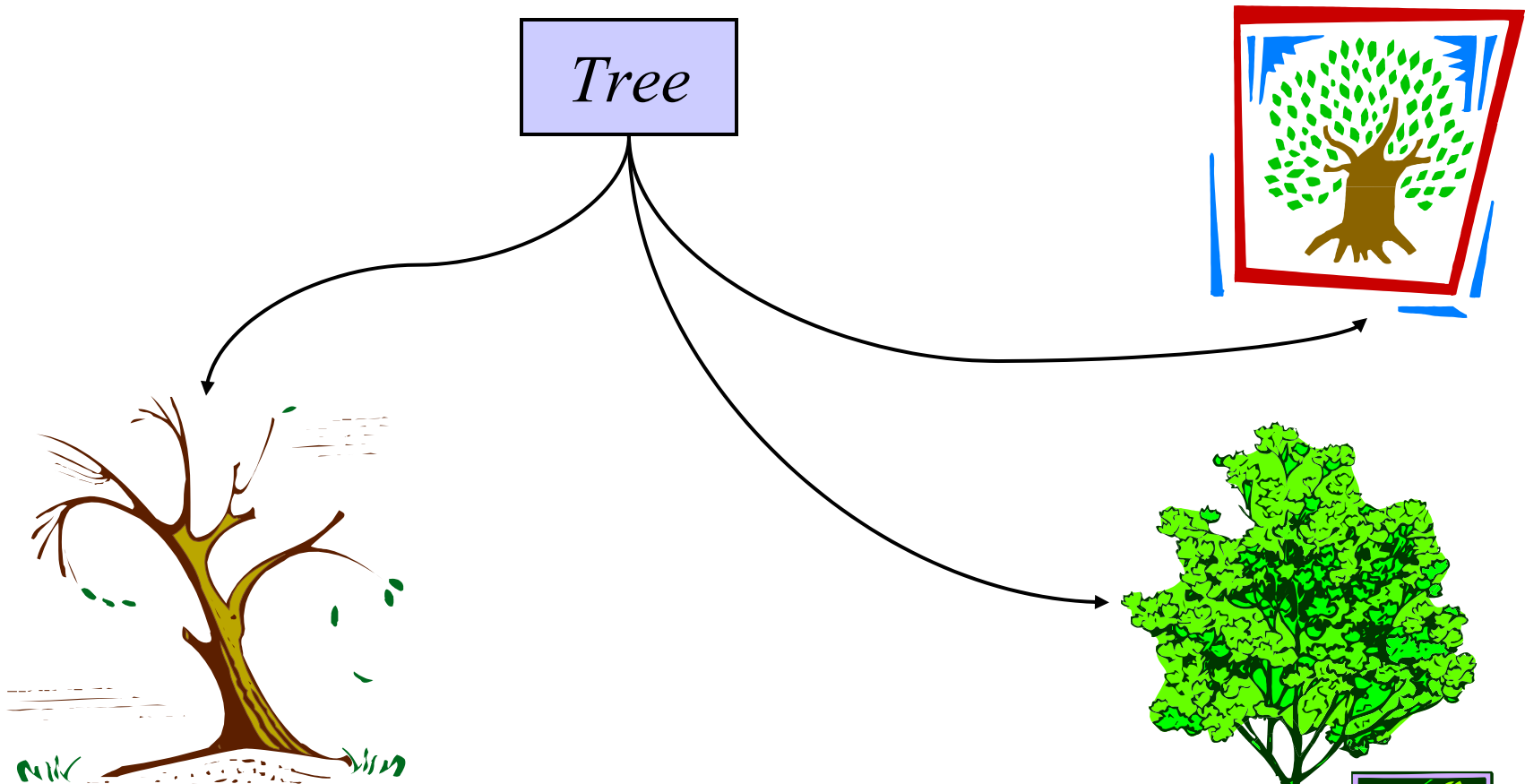
# Agent Issue #1:
# What Color is Blue?

Different languages and different cultures
do not map lexemes to the same referents.

*How do we know what standard is being used?*

# From Lexeme to Referent; A One-to-Many Map



*Tree*

# The Scope of a Lexeme

- To have meaning, every lexeme must be traceable to some real world reference(s).

- These referents differ in scope among:
    - <u>Languages</u>,
    - <u>Cultures</u>, and
    - <u>Individuals</u>.

# Language Scope

- Different Languages express "common concepts" with differing precision and referents
- Physics:
  - Color requires three dimensions in physics (e.g. Hue, Saturation, Intensity)
- Language
  - In English, blue is a single color;
    - In Russian it is two.
  - In Welsh, blue is "glas"
    - But it is also the color of growing things
  - Black, white, and grey have no hue;
    - are they colors?

# Cultural Scope

- <u>Culture</u>: Lexemes are defined in terms of cultural standards

- E.g., What makes someone "good" is often <u>determined by a comparison to ideal actions</u> in a process context.

  - 'A good person' will tell the truth in every circumstance.

# Individual Scope

Blue:

the color of the <u>clear sky</u>
<u>in the day</u>
viewed <u>from</u> the earth

## *What if you are color blind?*

# Unicorns

- What if there is no referent?
  - A camel is imagined beast in Antarctica

- We can construct new elements within lexical classifications by combining attributes without regard to "reality" or other constraints.

  - "A unicorn is the same as a horse except there is also long horn on its forehead …"

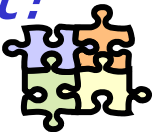  - We need these capabilities to express conjecture – ('what if …")

# Agent Issue #2: Hogs v. Pigs

You are selling hogs.
I am buying pigs.

*How do we know if*
*    my definition of a pig*
*    is sufficiently the same as*
*    your definition of a hog*
*    for the purpose of this transaction?*

*Would a steer be an acceptable replacement?*

# Natural Language can be ambiguous

- ## Because of <u>Syntax</u>
  - ### Pronouns
    - John and Sam were dueling He shot him.
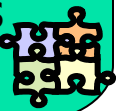  - ### Modifier Order
    - The man was dancing with a wooden leg.
- ## Because of <u>Overloading</u>
  - ### Words
  - ### Functions

> Ambiguity is unintentional;
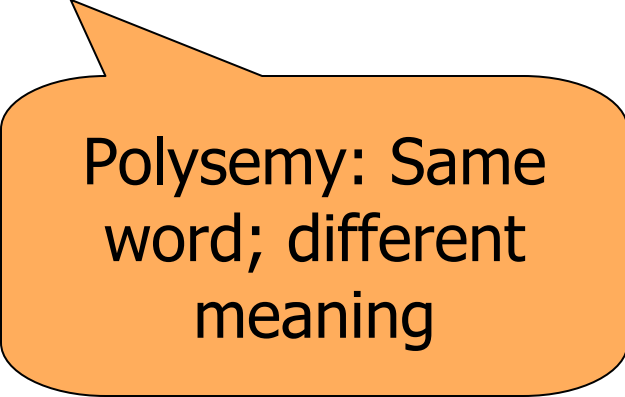> --------------------
> Generality is purposeful

# Hot ! - An overloaded word

- Spicy
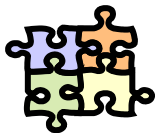- Warm
- Color
- Attractive
- Stolen
- Doing very well

Polysemy: Same word; different meaning

# Draft Requirements -- Referent

- No ambiguity !
  - No pronouns
  - No polysemy
  - No homonyms
  - No overloading of cases

# The Scope of a Lexeme

- To have meaning, every lexeme must be traceable to some real world reference(s).

- These <u>referents may differ in scope</u> among:
    - Languages,
    - Cultures, and
    - Individuals.

# Language Scope

- Different Languages express "common concepts" with differing precision and partitions in their referents

- Physics:
  - Color requires three dimensions in physics (e.g. Hue, Saturation, Intensity)

  See David Crystal

- But in Natural Language …
  - In English, blue is a single color;
    - In Russian it is two ("sinij" and "goluboj").
  - In Welsh, blue is "glas"
    - But glas is also the color of growing things
  - In Hanunoo, there are only four colors
  - Black, white, and grey have no hue;
    - Are they colors?

Copyright 2003, Everett M. Sherwood, Motorola

# Cultural Scope

- <u>Culture</u>: Lexemes are defined in terms of cultural standards

- E.g., What makes someone "good" is often <u>determined by a comparison to ideal actions</u> in a process context.
    - 'A good person' will tell the truth in every circumstance.

Copyright 2003, Everett M. Sherwood, Motorola

# Individual Scope

**Blue**:

> the color of the <u>clear sky</u>
> <u>in the day</u>
> viewed <u>from</u> the earth

## *What if you are color blind?*

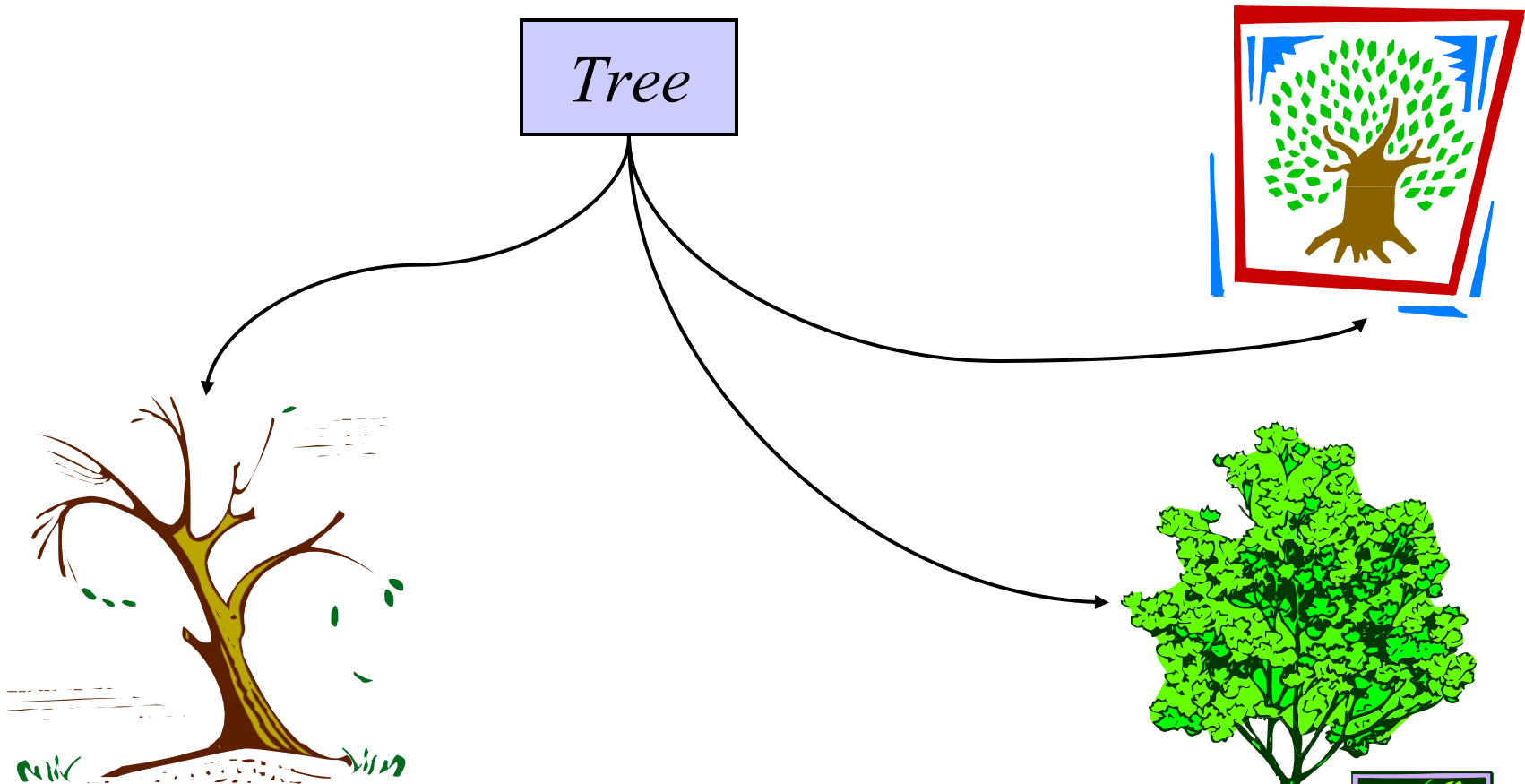# Unicorns

- What if there is no referent?
    - A camel is imagined beast in Antarctica

- We can construct new elements within lexical classifications by combining attributes without regard to "reality" or other constraints.

    - "A unicorn is the same as a horse except there is also long horn on its forehead …"

    - We need these capabilities to express conjecture – ('what if …")
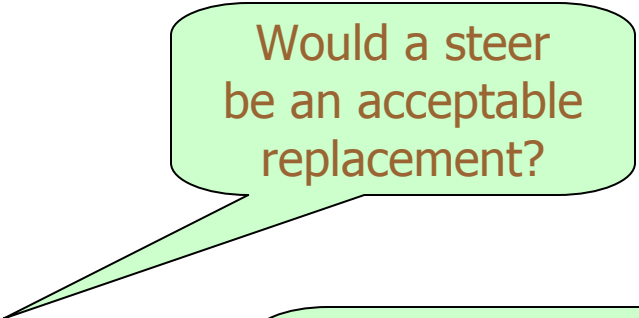
# From Lexeme to Referent;
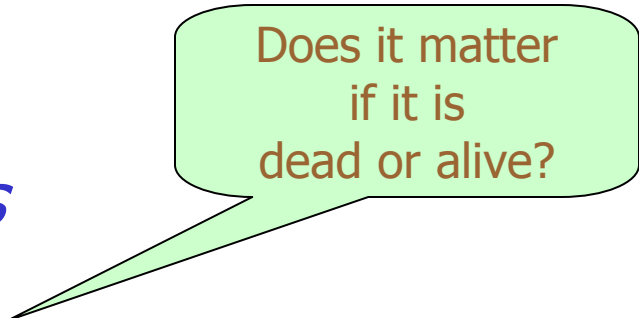# A One-to-Many Map



Tree

# Agent Issue #2: Hogs v. Pigs
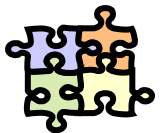
You are selling hogs.

I am buying pigs.

*How do we know if
my definition of a pig
is sufficiently the same as
your definition of a hog
for the purpose of this transaction?*

Would a steer
be an acceptable
replacement?

Does it matter
if it is
dead or alive?

# Natural Language can be ambiguous

- **Because of <u>Syntax</u>**
  - Pronouns
    - John and Sam were dueling *He shot him*.
  - Modifier Order
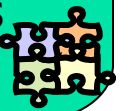    - The man was dancing <u>with a wooden leg</u>.
- **Because of <u>Overloading</u>**
  - Words
  - Functions

Ambiguity is unintentional;
-------------------
Generality is purposeful
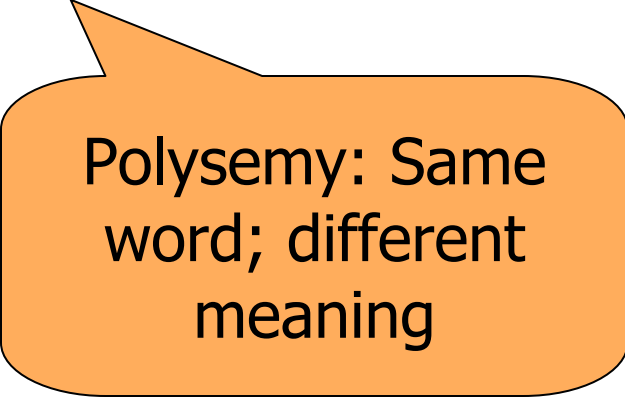
# Hot ! - An overloaded word

- Spicy
- Warm
- Color
- Attractive
- Stolen
- Doing very well

Polysemy: Same word; different meaning

# Plurals –

- Plural forms are a change in morphology to indicate <u>More Than One</u>

- More than one?
  - In English, the plural refers to more than one
  - In other languages, it can refer to more than two, or more than three …

- Polysemy Plurals – the plural changes the definition
  - Singular:
    - "Provision" – part of an agreement
  - Plural
    - "Provisions" – supplies
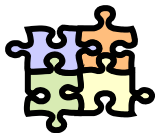
See David Crystal

# Closed vs. Open Word Classes

- Open Class
  - We make up new words
    - Nouns, Verbs, Adjectives, Adverbs, …

- Closed Class
  - We use only the words already available:
    - Conjunctions
    - Demonstratives
    - Quantifiers
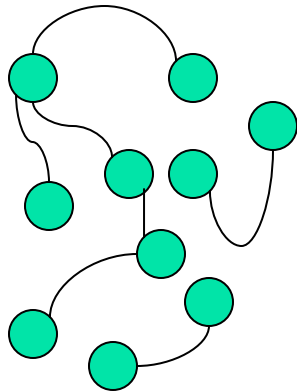    - Prepositions

# Draft Requirements – For Lexeme Referents

- No ambiguity !
  - No pronouns
  - No polysemy
  - No homonyms
  - No overloading of cases

# Lexical Structures

The relationships of words
to their referents
and with one another

# Overlapping Lexical Structures

- Dictionaries
  - OED – Oxford English dictionary

- Word Nets
  - Roget's Thesaurus
  - May or may not be included in a dictionary
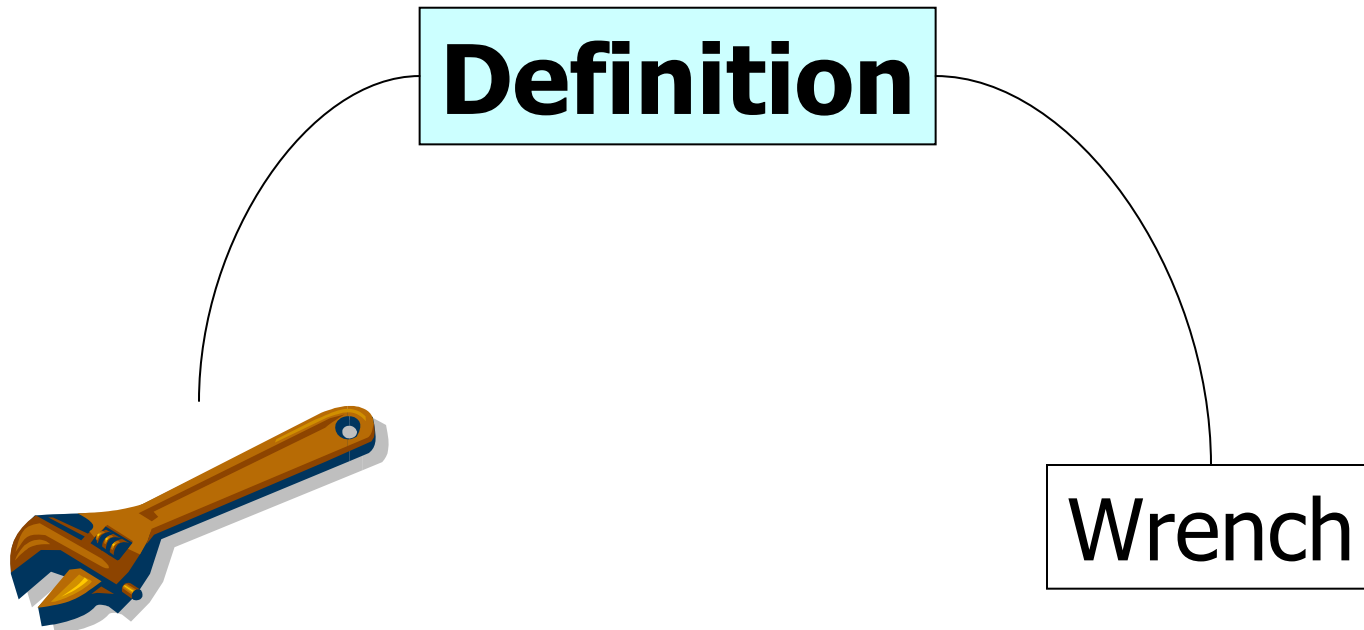
# Dictionaries

More than a difference

# The Lexicographer's Task

**Definition**

Wrench

# Agreement of words is the basis for communication

*What do <u>you</u> mean by tree?*

Katie's Tree Definition

Has Height     Has Color     Is Near Water

Jan's Tree Definition

Has Height     Has 1st Color     Provide fire     Has 2nd Color

The AGREED Definitions of Jan and Katie

**Dictionary**

# Types of Definitions

See George Miller

- ## Constructive
  Dictionary
  - Provides the information needed to construct an instance

- ## Differential
  Word Net
  - Assumes meaning of at least one related term is already understood
    - "Canine" may be defined by reference to: "dog", "wolf", and "fox"
  - Provides the information needed to confirm / distinguish an instance

# A Constructive Use
# Of A Differential Definition

- Dictionary definition of "Erode"
  - "To eat out, to eat away"
- A little girl wrote:
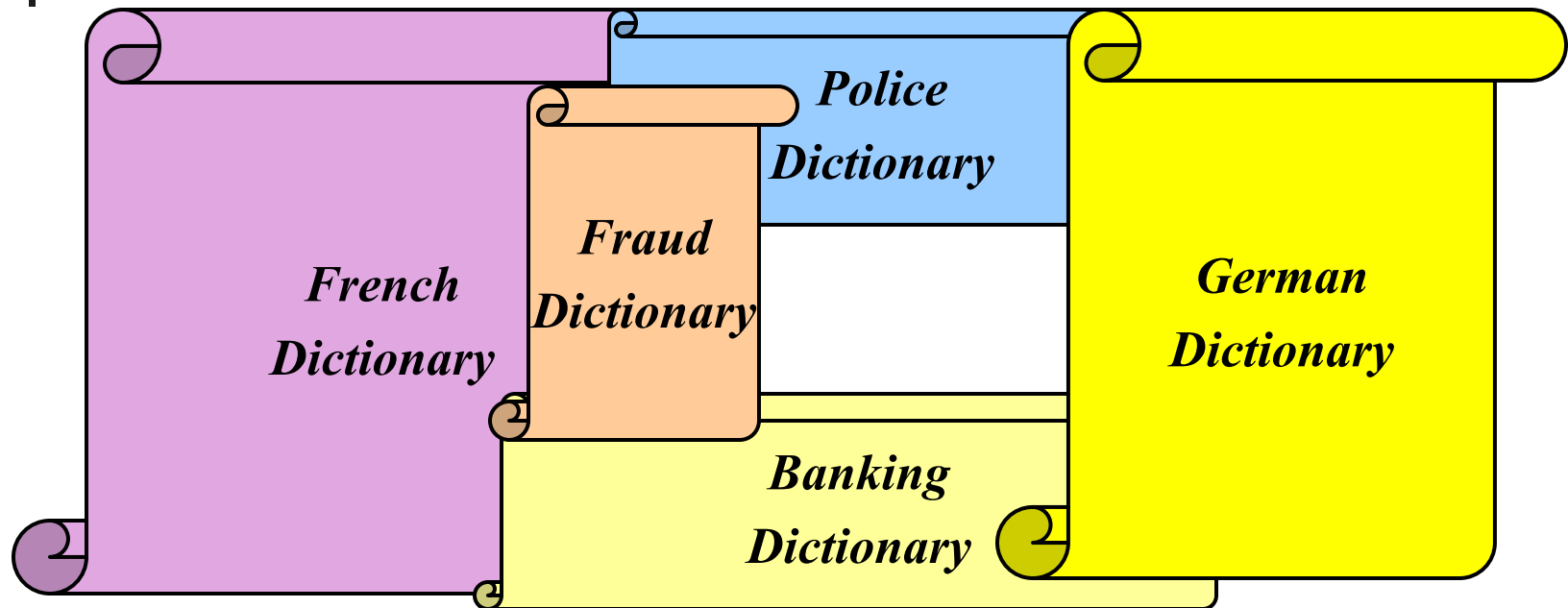
    *"Our family erodes a lot."*

    -- George Miller

# What Constitutes a Definition?

- ... Some Criteria
  - A function based on attributes and their values
- ... or A Prototype
  - An "average" of attributes
- ... or Exemplars
  - A list of examples
- ... or A Standard
  - An agreed referent
    - E.g. the standard meter in Paris

# There is more than one Dictionary.



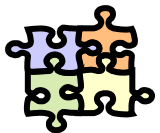*Agent Dictionaries will need to include Decision Functions.*

# A Word ... is pointer to a Description with an embedded Decision

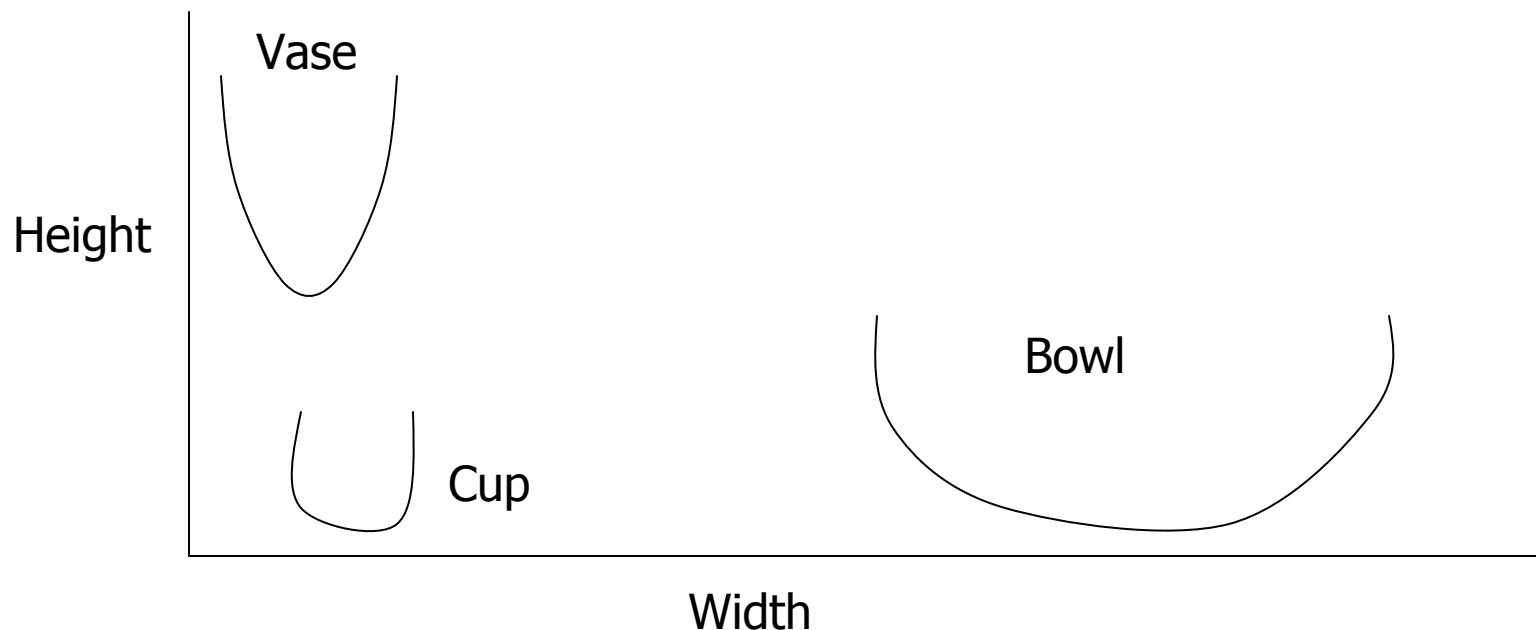When we talk about trees, we refer to:

- A description
  - Tall, green, plant, ...

- Plus ... the criteria to determine if an individual instance matches the description
  - A tree is larger than a shrub

# When is a cup a bowl?

- The criteria gives a different output
  as there is a change in the values of attributes
  (height, width)

# Different Criteria Produce Different Structures

See George Lakoff

- Cladists vs. Pheneticists -- Different Criteria for Biological Taxonomy
  - Pheneticists:
    - Similarity in form, function, and biological role
    - "If it walks like a duck and it talks like a duck …"
  - Cladists:
    - Shared, derived evolutionary characteristics

**"… I regret to report that there is surely no such thing as a fish."**
– Steven J. Gould
(What, If Anything is a Zebra? 1983)

Copyright

# Draft Requirements - Definitions

- We must establish the Scope of the Lexemes
  - The Referents subtended by a Lexeme
  - Its Affectivity: time; location; conditions …

- There must be One True Source of Definition
  - A single Location
    - A definition
      - External reference (human reference )
      - Internal
    - Datatype (image, sound, …)
  - A decision function
    - Tells what matches or does not match

- A responsible party must maintain the definition.
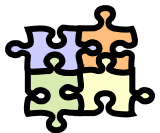
# Word Nets
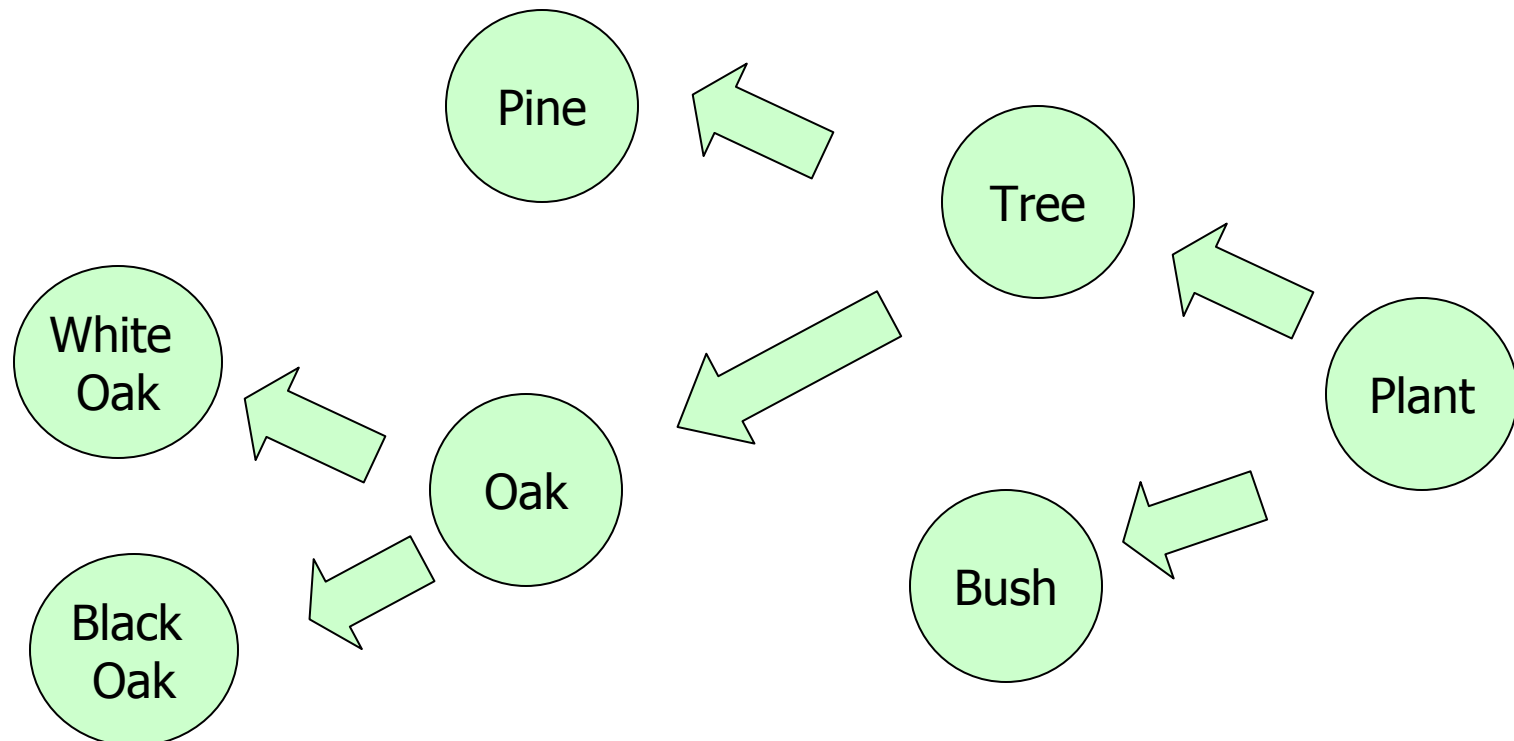
"You shall know a word
by the company it keeps."

-- J. R. Firth

# Suppose there was only one word ...

- We <u>construct</u> new words to portray <u>differences.</u>

Copyright 2003, Everett M. Sherwood, Motorola

# Differences Provide Meaning

1) Distinctions in word definitions

2) Differences in How Words Relate to One Another

*We use lexical structures to identify and to organize the differences among lexemes.*

A rich set of definitions and relations allow us to find "just the right word."

Copyright 2003, Everett M. Sherwood, Motorola

# Lexemes are related in Lexical Structures

- The lexical structure describes how words are semantically differ from one another.

- Words can be semantically associated by …
  - Synonymy
  - Antonymy
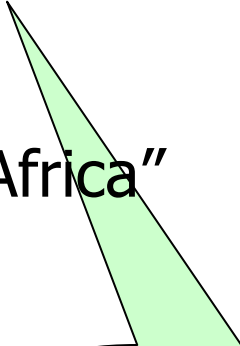  - Meronymy
  - Hyponymy
  - Incompatibility
  - Compatibility

These associations are not relationships; they are the types of relationships

# Synonymy

- Synonyms "mean the same" (almost)

- Lexical classifications can be very different from real-world classifications
  - Worms are regarded as insects (in Japanese)

- Synonyms depend on <u>conditions</u>!
  - "Dutch" may be a synonym for "South Africa"

Th. R Hofmann

# Antonymy

- Antonyms
  - One term excludes the others
    - Think: "radio buttons"
  - Gradable: an ordinal scale
    - High/Low  allows Very high, ... very low
  - Complements  - binary values only
    - On/ Off

- Antonymous groups
  - Mutually exclusive values
  - Selection list: Rose, Iris
  - Selection list: North, South, East, West
    - AND North is opposite to South ...
- Cyclical
  - Sunday, Monday, ... Saturday
- Hierarchies
  - Ranks in the Military

# Hyponymy

- Hyponym – a term whose meaning is included in another term
  - Dog is a hyponym of animal
  - Animal is a hypernym of dog

- Hyponyms are more restrictive than their hypernym
  - In a data model, they have additional attributes or limited values of attributes

- The restrictions may not imply alternative values
  - Lamb is a young sheep
  - There is no single word for "not young sheep"

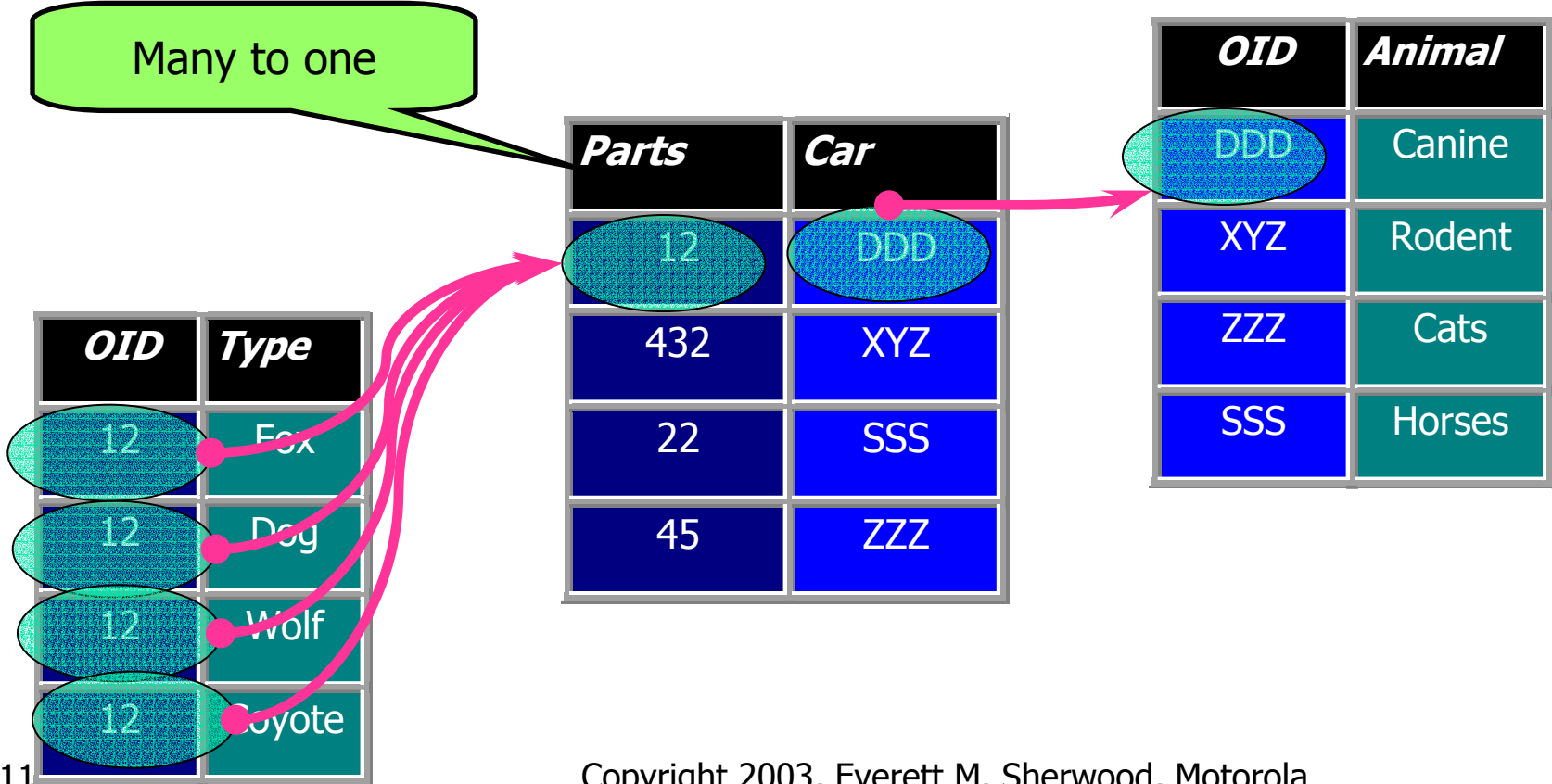- In data models, they occur in relationships called
  - "IS-A" or "A KIND OF"

Th. R Hofmann

# Hypo/Hypernyms -- "IS A"

*"A fox is a canine."*

Many to one

| OID | Type |
|-----|------|
| 12  | Fox  |
| 12  | Dog  |
| 12  | Wolf |
| 12  | Coyote |

| Parts | Car |
|-------|-----|
| 12    | DDD |
| 432   | XYZ |
| 22    | SSS |
| 45    | ZZZ |

| OID | Animal |
|-----|--------|
| DDD | Canine |
| XYZ | Rodent |
| ZZZ | Cats   |
| SSS | Horses |

Copyright 2003, Everett M. Sherwood, Motorola

# Lexical Blocking -- a side effect of hyponymy

Th. R Hofmann

To convey precision in our descriptions,
  we choose words that are neither too general nor too specific.

- There is a cow in my office.
- There is an animal in my office.
  (Too general: Conveys insufficient information)
- There is a Holstein in my office.
  (Too specific: Conveys useless information)

- To convey meaning, the appropriate term "blocks" the others.

- Blocking may be partial
  - Is a thumb a finger?

- The boundary may be vague
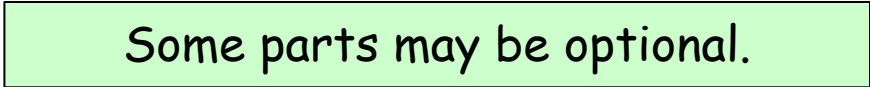  - Is a hill a kind of mountain?

# The Manner of ... Troponyms

- **Hyponyms for verbs**
  - E.g., Stagger is a troponym of walk

# Part / Whole  -- Meronymy

- Think: "Bill of Materials"

- Whole/part relationships are different than hyponyms
  - A leg is part of a chair but not a kind of a chair.
  - Chair is NOT a hypernym of leg.

- In data models these relationships are called "HAS-A"

Some parts may be optional.

Holonym –Part/Whole

# Part / Whole -- "HAS A"

*"The green Mercedes has a wheel and has a handle …."*

One to Many

| Parts | Car |
|-------|-----|
| 12 | DDD |
| 432 | XYZ |
| 22 | SSS |
| 45 | ZZZ |

| OID | Part | Cost |
|-----|------|------|
| 12 | Wheel | 34 |
| 12 | Handle | 5 |
| 22 | Computer | 288 |
| 45 | Knob | 2 |

| OID | Car | Color |
|-----|-----|-------|
| DDD | Mercedes | Green |
| XYZ | Toyota | Blue |
| ZZZ | Ford | Red |
| SSS | Volvo | Red |

# Types of Meronymy

David Crystal

- Component/Object (branch/tree)
- Member/Collection (tree/forest)
- Portion/Mass (slice/cake)
- Feature/Activity (paying/shopping)
- Place/Area (Idaho, USA)
- Phase/Process (Child/Person)

# Constraints: Semantic Incompatibility

- Incompatible – a term whose sense is <u>excluded</u> by another
  - A flower that is a wind

- Sometimes the semantics are unclear
  - Should "black, white, and gray" be allowed as colors?

# Semantic Compatibility

George Miller

- Words that <u>may</u> 'go together"
- Mother
  - Birth, Foster, …
  - Black
    - But not green

# Lexical Claims

- There will be alternate lexical structures
- Remember that different sources will produce different structures
  - Lexical $\neq$ Lexical
  - Lexical $\neq$ Physical
  - Physical $\neq$ Physical
- These differences occur not only across languages but also within languages
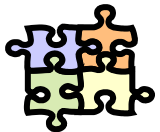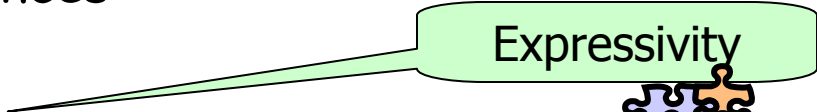
# Summary (so far)

- Linguistics is a source of semantic requirements
- Low level semantics is the basis for "making sense."
- "Making sense" is the first requirement of Agent Semantics
- Making Sense depends on:
    - Mapping words with their referents
    - Combining those words into sentences

    Representation

- Better Sense depends on:
    - Choosing the right words based on lexical relation

    Expressivity
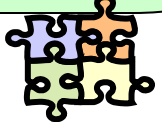
# Next -- Part 2.  Sentences

A sentence is a semantic unit
that is built by combining words with one another.

Of interest is how the words may be combined.