

Generative Lexicon Theory: Integrating Theoretical and Empirical Methods

James Pustejovsky
Brandeis University

Elisabetta Ježek
University of Pavia

July 11-15, 2016
NASSLLI 2016
Rutgers University



- July 11: Introduction to GL and Data Analytics
- July 12: Qualia Structure
- July 13: Event Structure
- July 14: Argument Structure
- July 15: Meaning Composition

Introduction to Generative Lexicon

- Basic concepts in GL
 - Motivation
 - Notation and Language: typed feature structures
 - Meaning Composition in GL
- Polysemy and the Lexicon-Pragmatics Interface
- Evidence-based linguistics and data analytics

Qualia Structure

- What is a Quale?
- What motivates Qualia?
- Default Qualia and context updating
- Methodology to identify Qualia
- Data for each Quale
- Qualia and Conventionalized Attributes
- Qualia for Verbs

Lab on Qualia identification and extraction

Event Structure

- Events as Structured Objects
- Event Types
 - States
 - Transitions
 - Point Verbs
 - Processes
- Events as Labeled Transition Systems
- Dynamic Event Models

Lab on identification of event types

Argument Structure

- Argument Types in GL
 - True Arguments
 - Shadow Arguments
 - Hidden Arguments
- Argument Structure Representation
- Arguments and Defaulting

Lab on hidden and shadow arguments

Meaning composition

- Basic Assumptions
- Simple Function Application
- Coercion
- Subselection
- Co-composition

Lab or assignment on coercion

Lecture 2: Qualia Structure

- What is a Quale?
- What **motivates** Qualia?
- **Default** Qualia and context **updating**.
- **Methodology** to identify Qualia
- **Data** for each Quale.
- Qualia for **Verbs**.

Causation as Understanding

- Hume's cause: Counterfactual future (Agentive quale)
- Projective cause: Telic quale
- Inherent cause: Constitutive quale
- Class distinction: Formal quale

Moravcsik's 1975 interpretation of Aristotle's notion of *Aitia*

What is a *Quale*?

- A *Quale* (singular of *Qualia*) indicates a **single aspect of a word's meaning**, defined on the basis of the **relation** between the **concept** expressed by the word and **another concept** that the word evokes.
- Among the conceptual relations that a word may activate *Qualia* relations as defined in GL are those that are **exploited in our understanding of linguistic expressions**.
- *fresh bread* = “bread which has been **baked** recently.”

Contextual modulations of noun meaning, due to the selecting predicate.

- This car weighs over 2,000 lbs.
- We buy vehicles such as cars and buses.
- John started the car.
- You should warm your car up in winter.
- Did you lock the car?
- The car screeched down the road.

Linguistic phenomena motivating Qualia relations

Inference of implicit predicates from particular constructions.

- Verb-Noun combination: finish the beer/cake.
- Adjective-Noun combination: comfortable chair/shoes.
- Noun-Noun compounding: dinner dress/table.

Light verbs specifications.

- take a tablet.
- take a train.

Noun-to-Verb transformations.

- fax a document.
- microwave the chicken.
- lace the shoes.

Linguistic phenomena motivating Qualia relations

Short passives.

- *This picture was painted.
- This picture was painted in 1604.

Middles.

- *This book reads.
- This book reads easily.

Adjectival Use of Past Participles.

- *a built house.
- a recently built house.

Decomposition in GL

- Traditional decomposition frameworks fail to account for the **semantic flexibility of words** shown in the previous examples.
- GL draws insights about the meaning of a word by looking at **the range of its contextual interpretations**, and by examining how this range can be predictably derived from the underlying meanings.
- On this basis, the meaning is decomposed into the relevant features.

Qualia Relations or Roles

- **Formal** (F): encoding taxonomic information about the lexical item (the *is-a* relation);
- **Constitutive** (C): encoding information on the parts and constitution of an object (*part-of* or *made-of* relation);
- **Telic** (T): encoding information on purpose and function (the *used-for* or *functions-as* relation);
- **Agentive** (T): encoding information about the origin of the object (the *created-by* relation).

Qualia Structure

$$\left[\begin{array}{l} \alpha \\ \text{QUALIA} = \left[\begin{array}{l} \text{F} = \text{what } x \text{ is} \\ \text{C} = \text{what } x \text{ is made of} \\ \text{T} = \text{function of } x \\ \text{A} = \text{origin of } x \end{array} \right] \end{array} \right]$$

Qualia Structure

$$\left[\begin{array}{l} \text{car} \\ \\ \text{QUALIA} = \left[\begin{array}{l} \text{F} = \text{vehicle} \\ \text{C} = \text{engine, door, wheels,...} \\ \text{T} = \text{drive} \\ \text{A} = \text{build} \end{array} \right] \end{array} \right]$$

Methodology for identifying Qualia Values

- **Linguistic evidence** determines what information is stated to be lexically associated with the Qualia Structure of a word.
- If a piece of knowledge is **regularly exploited** in our understanding of linguistic expressions, then it is likely to be part of lexical information and it is included in QS as **default value** for a Quale.
- Pustejovsky and Jezek 2012.

Identifying Qualia Values

- The rock shattered the window. ([C = pane])
- Wooden windows are prone to rotting. ([C = frame])
- John was going to paint his room. ([C = walls])
- She has swept the room. ([C = floor])
- The knife cut his finger. ([C = blade])
- sharp knife. ([C = blade])

Identifying Qualia Values

- Do you want the whole house waken up? ([T = live_in(human, building)])
- The rest of the house was sleeping. ([T = live_in(human, building)])
- They crossed the river. ([F = space])
- The river is wide. ([F = space])
- The river had frozen during the severe weather. ([C = water])
- The river became polluted. ([C = water])
- the banks of a polluted river. ([C = water])

- Not all lexical items carry a value for each Q-role.
- Some values are **left unspecified**, while others are populated with **more than one value**.
- Nouns denoting natural kinds (e.g., *rock*, *fish*, *air*, *sea*) typically do not have a value for the A.

Default values and context updating

- Lexical meaning provides **default values** for the different Q-roles.
- ... an insect as small/black as an ant...
- ?a small/black ant.
- Default values may be **updated** from discourse context in composition.
- a large ant.
- Unspecified Q-values may be **introduced** in context.
- ?black ant.
- black dog.

The Formal

- The Formal (F) encodes the relation between the **entity** denoted by the word and the **category** it belongs to.
- This relation enables one to grasp the nature of an entity by **discriminating it from other kinds**.
- What **type of entity** is x denoting? *rock* denotes a natural kind, *table* denotes an artifact, *car* denotes a vehicle, *park* denotes a location, *water* denotes a liquid, *plant* denotes a living thing, *fish* denotes an animal, *hand* denotes a body part, *glass* denotes a container, and so on.

- **More classifications** are possible for the same type of object: for example, a *knife* can denote both a weapon or a kitchenware.
- Classifications **at different levels of generalization** are available for reference:
 - a liquid such as water.
 - fluids such as water or air.
 - substances such as fluids, salts, glucose and carbon dioxide.

Formal factors for the class of nouns denoting concrete entities

- Spatial characteristics, intrinsic orientation.
- Size and dimensional properties.
- Shape and form.
- Color.
- Position.
- Surface.

Values for Formal factors of nouns denoting concrete entities

- a red car (Color_F)
- a long dress (Dimension_F)
- a round table (Shape_F)
- a red pen (Color_F) or T/C (depending on contextual interpretation)
- a flat screen (Shape_F)
- a thick sweater (Dimension_F)

Values for Formal factors of nouns denoting concrete entities

- the length of the table (Dimension_F)
- the facade of the building (Orientation_F)
- wipe the floor (Surface_F)
- a large (Size_F) round (Shape_F) table
- *a round (Shape_F) and square (Shape_F) table

Formal-specific Constructions

- *NP such as NP*: *events such as lectures, walks, tours and meetings*;
- *such NP as NP*: *such areas as children's playground*;
- *NP and other NP*: *rum and other spirits*;
- *NP or other NP*: *insects or other animals*
- *NP, including NP*: *recyclable materials including glass*;
- *NP, especially NP*: *cool temperate countries especially Europe and North America*;
- *favorite NP is NP*: *Mario's favorite food is pasta*.

Multiple inheritance over the Formal Quale

- Some lexical items appear to have more than one F value (**multiple inheritance** over F).
- Hand me the book. (book as physical_object)
- The author discussed her book. (book as information)
- It is a bulky and demanding book. (book as phys•info)
- In the GL model, these are called **dot object nominals** (Cruse's 2004 **nouns with facets**, Asher's 2011 **dual aspect nouns**).

The Constitutive Quale

- The Constitutive (C) role encodes information about what is “inside” the object denoted by the word, particularly the material the object is **made of** (i.e., its stuff), and the **parts it consists of**.
- There is a fundamental distinction between **inherently individuated things**, such as humans, tigers, and trees, and **inherently undifferentiated stuff**, such as water, air, and sand.
- This is reflected in the **count** vs. **mass** distinction in linguistics.

Count Nouns and Mass nouns in GL

COUNT NOUN: (where $\alpha \neq \beta$).

$$\left[\begin{array}{l} \mathbf{N} \\ \text{QUALIA} = \left[\begin{array}{l} \mathbf{F} = \alpha \\ \mathbf{C} = \beta \end{array} \right] \end{array} \right]$$

MASS NOUN:

$$\left[\begin{array}{l} \mathbf{N} \\ \text{QUALIA} = \left[\begin{array}{l} \mathbf{F} = \alpha \\ \mathbf{C} = \alpha \end{array} \right] \end{array} \right]$$

The Formal-Constitutive Equivalence Constraint for mass nouns in GL

$$\left[\begin{array}{l} \textit{water} \\ \text{QUALIA} = \left[\text{F/C} = \textbf{liquid} \right] \end{array} \right]$$
$$\left[\begin{array}{l} \textit{rock} \\ \text{QUALIA} = \left[\text{F/C} = \textbf{solid_substance} \right] \end{array} \right]$$

but cf.

$$\left[\begin{array}{l} \textit{rock} \\ \text{QUALIA} = \left[\begin{array}{l} \text{F} = \textbf{solid_object} \\ \text{C} = \textbf{solid_substance} \end{array} \right] \end{array} \right]$$

Default values for MADE-OF relation in C

- The river had frozen during the severe weather. ([C = water])
- The river became polluted. ([C = water])
- the banks of a polluted river. ([C = water])

$$\left[\begin{array}{l} \textit{river} \\ \text{QUALIA} = \left[\begin{array}{l} \text{F} = \text{space} \\ \text{C} = \text{water} \end{array} \right] \end{array} \right]$$

Nominal compounding.

- plastic bag
- paper cup
- leather shoes
- gold watch
- milk chocolate

Adjective-Noun constructions.

- a golden ring
- a wooden floor
- a metallic paint

The Inverse Constitutive C_I introduces the relation between an object and the larger object it is logically a part of, if such a dependency exists.

$$\left[\begin{array}{l} \text{roof} \\ \text{QUALIA} = \left[\begin{array}{l} F = \text{phys} \\ C_I = \text{building} \end{array} \right] \end{array} \right]$$

Constitutive-specific constructions

- N_1 's N_2 : $C(N_2, N_1)$ *the room's wall.*
- N_2 of N_1 : $C(N_2, N_1)$ *the door of the car.*
- NP_2 is a part of NP_1 : $C(NP_2, NP_1)$ *brain is a very sensitive part of the body.*
- NP_1 made of NP_2 : $C(NP_2, NP_1)$ *monuments made of stone and marble.*
- NP_1 of NP_2 : $C(N_2, N_1)$ *house of wood.*
- NP_1 consists of NP_2 : $C(NP_2, NP_1)$ *the orchestra consists of ninety performers.*
- NP_1 containing NP_2 : $C(NP_2, NP_1)$ *a forest containing dead trees.*

Hidden events and the Telic

- The Telic relation (T) encodes information about the **intended use or function** of an object.
- It expresses the relation that allows us to grasp what an entity is by **knowing what it is used for**.
- It encodes a **potential activity** of the object.
- First systematic mention of Telic in Pustejovsky and Anick (1988) as **hidden event**.

Default values for Telic relation

- Any chocolate? Not after that cake! ([T = eat])
- I prefer cake to biscuits. ([T = eat])
- We skipped the cake and settled for another coffee. ([T = eat])
- the next customer. ([T = take_care_of])
- the next slide. ([T = project])
- This is a difficult problem. ([T = solve])
- This is a difficult question. ([T = solve])

Updating Telic values in composition

When the Telic activity being expressed does not correspond to the Telic value specified in the noun, we say that the expression updates the Telic information associated with the noun in composition.

$$\left[\begin{array}{l} \textit{shopping bag} \\ \text{QUALIA} = \left[\begin{array}{l} \text{F} = \textbf{container} \\ \text{T} = \textbf{shopping} \end{array} \right] \end{array} \right]$$

Telic for different types and roles

- functional locations: *library, gym, church, school.*
- professions: *doctor, teacher, lawyer.*
- agentive nominals (individuals engaged in an activity, either habitually or occasionally): *runner, passenger, movie goer.*

Telic-specific constructions

- an NP to V: $T(V, NP)$ *a book to read.*
- an NP worth V-ing: $T(V, NP)$ *a question worth asking.*
- the NP merits/deserves V-ing: $T(V, NP)$ *This book deserves reading..*
- enjoy/prefer V-ing NP: $T(V, NP)$ *enjoy listening to music / prefer watching television.*
- an Adj NP to V: $T(V, NP)$ *a difficult question to ask.*
- an NP (used) for V-ing: $T(V, NP)$ *a spade (used) for digging.*

- The Telic of a natural kind (T_N), (*human, dog, water, and so forth*) encodes information about the **actions and properties that the object engages in**, but that are not intentional or purposive.

A *river* does not intentionally flow, but this is a **necessary property of a body of water** if it is to qualify as a river (as in a fast / rapid / slow / lazy river).

$$\left[\begin{array}{l} \textit{river} \\ \text{QUALIA} = \left[\begin{array}{l} \text{F} = \text{space} \\ \text{C} = \text{water} \\ \text{T}_N = \text{flow} \end{array} \right] \end{array} \right]$$

- The Agentive quale (A) encodes information about the origin of the object denoted by N.
- It provides a mechanism for discriminating those objects that present themselves to us (occurring **naturally**) from the various **artifacts** that we create through our own activities and intentional behavior.
- Differently from the value of T, A introduces an **existentially bound** or **existentially quantified** event, that precedes the existence of the object.

Agentive-Telic Pairing

- Inherent in this is an **association** between the Agentive and Telic of the object, i.e., the object is made for a purpose (Agentive-Telic pairing).
- Natural kinds lack this association, as they do not encode an Agentive value.

Default values for Agentive

- He just finished and published his first novel. ([A = write])
- Woody Allen has started a new movie. ([A = direct, film])
- John began a large oil painting yesterday. ([A = paint])
- Mary made a cake. ([A = bake])
- Her mother made her a dress. ([A = sew])
- fresh coffee ([A = brew])
- fresh water (in contrast to “salt water”)

Artifactual Types with an Agentive (coffee)

$$\left[\begin{array}{l} \textit{water/coffee} \\ \text{QUALIA} = \left[\begin{array}{l} \text{F} = \textbf{liquid} \\ \text{T} = \textbf{drink} \end{array} \right] \end{array} \right]$$

$$\left[\begin{array}{l} \textit{coffee} \\ \text{QUALIA} = \left[\begin{array}{l} \text{F} = \textbf{liquid} \\ \text{T} = \textbf{drink} \\ \text{A} = \textbf{brew} \end{array} \right] \end{array} \right]$$

CA and conventionalized commonsense knowledge

- A conventionalized attribute (CA) is a property typically associated with an object through our experiencing of it.
- Only **the most conventional activities** associated with an object are coded in the noun's meaning as CAs.
- They can be identified through **empirical testing**.

Conventionalized Attributes

- They heard the village dog in the distance. ([SOUND = barking])
- Ann was listening to the birds. ([SOUND = singing])
- He could hear the rain in the garden. ([SOUND = falling])
- John can smell the flowers in his garden. ([SMELL = scent])
- The repairman smelled gas in the kitchen. ([SMELL = odor])
- Mary woke up and smelled coffee. ([SMELL = aroma])

Conventionalized Attributes

- Mary sat out and enjoyed the sun. ([CA = warming up])
- It's a great place to enjoy the sea. ([CA = viewing, swimming, walking])
- The tuna is one of the fastest fish in the sea. ([CA = swimming])
- John was the fastest boy in the school. ([CA = running])
- I could hear a car behind me. ([CA = driving])
- We do occasionally hear an airplane. ([CA = flying])

CAs as projective manifestations of Qualia

- CAs in GL are interpreted as projective manifestations of specific Qualia.
- The representation of CAs in GL is always mediated through a Quale.

Projective operations over T values

- There's no train till 7:00 pm. (there is no departing)
- The train was delayed for an hour. (the departure was delayed)
- I left in time to catch the early train. (departing early)
- Heavy foods such as dairy products and meat.

Representation of Conventionalized Attributes

$$\left[\begin{array}{l} \textit{dog} \\ \text{QUALIA} = \left[\text{F} = \left[\begin{array}{l} \text{F} = \textbf{animal} \\ \text{CA} = \textbf{bark} \end{array} \right] \right] \end{array} \right]$$

$$\left[\begin{array}{l} \textit{fish} \\ \text{QUALIA} = \left[\text{F} = \left[\begin{array}{l} \text{F} = \textbf{animal} \\ \text{CA} = \textbf{live_in(water), swim} \end{array} \right] \right] \end{array} \right]$$

$$\left[\begin{array}{l} \textit{car} \\ \text{QUALIA} = \left[\begin{array}{l} \text{F} = \textbf{vehicle} \\ \text{T} = \left[\begin{array}{l} \text{T} = \textbf{drive} \\ \text{CA} = \textbf{make_noise} \end{array} \right] \end{array} \right] \end{array} \right]$$

Qualia Structure for Verbs

- When applied to event-denoting expressions, QS receives the following different interpretation in the model.
- F characterizes predicates denoting stable and persistent verbal predicates, namely states such as *love* and *believe*.
- A refers to the manner in which something **happens** or **changes**, i.e. it introduces the causing act or the process of verbs such as *run* and *walk*.
- T introduces the purpose of actions with verbs denoting intentional acts, such as *build* and *clean*.

F can be seen as characterizing predicates denoting stable and persistent verbal predicates, namely states such as *love* and *believe*.

STATE:

$$\left[\begin{array}{l} \textit{love} \\ \text{QUALIA} = \left[\text{F} = \mathbf{\textit{love_state}} \right] \end{array} \right]$$

Agentive for Verbs

Intentional activities such as those denoted by the verbs *run* and *walk* can be characterized as Agentive Quale verbs.

ACTIVITY:

$$\left[\begin{array}{l} \textit{run} \\ \text{QUALIA} = \left[\text{A} = \text{run_act} \right] \end{array} \right]$$

Qualia for Change-of-State verbs

Change-of-state verbs such as *break* and *open* can be modeled as denoting a static resulting state (Formal) brought about by an activity (Agentive).

CHANGE_STATE:

$$\left[\begin{array}{l} \textit{break} \\ \text{QUALIA} = \left[\begin{array}{l} \text{F} = \mathbf{broken} \\ \text{A} = \mathbf{break_act} \end{array} \right] \end{array} \right]$$

Intentional or directed events such as *build* and *clean* can be viewed as denoting a static intended goal state (Telic) brought about by an activity (Agentive).

ACCOMPLISHMENT:

$$\left[\begin{array}{l} \textit{build} \\ \text{QUALIA} = \left[\begin{array}{l} \text{T} = \text{build_goal} \\ \text{A} = \text{build_act} \end{array} \right] \end{array} \right]$$

GL Feature Structure

$$\left[\begin{array}{l} \alpha \\ \text{ARGSTR} = \left[\begin{array}{l} \text{ARG1} = x \\ \dots \end{array} \right] \\ \text{EVENTSTR} = \left[\begin{array}{l} \text{EVENT1} = e_1 \\ \text{EVENT2} = e_2 \end{array} \right] \\ \text{QUALIA} = \left[\begin{array}{l} \text{CONST} = \text{what } x \text{ is made of} \\ \text{FORMAL} = \text{what } x \text{ is} \\ \text{TELIC} = e_2: \text{function of } x \\ \text{AGENTIVE} = e_1: \text{how } x \text{ came into being} \end{array} \right] \end{array} \right]$$

Type Composition Logic (Asher and Pustejovsky, 2006)

1. e the general type of entities; t the type of truth values.
(σ, τ range over all simple types, and subtypes of e .)
2. If σ and τ are types, then so is $\sigma \rightarrow \tau$.
3. If σ and τ are types, then so is $\sigma \otimes_R \tau$; R ranges over A or T .
4. If σ and τ are types, then so is $\sigma \bullet \tau$.

Qualia Types

$$\left[\begin{array}{l} x: \quad \alpha \\ \quad \otimes_c \beta \\ \quad \otimes_t \tau \\ \quad \otimes_a \sigma \end{array} \right]$$

with an unlabeled qualia value

$$\left[\begin{array}{l} x: \quad \alpha \\ \quad \otimes \tau \end{array} \right]$$

Entities formed from the application of the FORMAL and/or CONST qualia roles:

1. For the predicates below, e_N is structured as a join semi-lattice, $\langle e_N, \Xi \rangle$;
2. *physical, human, stick, lion, pebble*
3. *water, sky, rock*

Natural Predicate Types

Predicates formed with Natural Entities as arguments:

1. *fall*: $e_N \rightarrow t$
2. *touch*: $e_N \rightarrow (e_N \rightarrow t)$
3. *be under*: $e_N \rightarrow (e_N \rightarrow t)$
 - a. $\lambda x: e_N [\textit{fall}(x)]$
 - b. $\lambda y: e_N \lambda x: e_N [\textit{touch}(x,y)]$
 - c. $\lambda y: e_N \lambda x: e_N [\textit{be-under}(x,y)]$

Artifactual Entity Types

Entities formed from the Naturals by adding the AGENTIVE or TELIC qualia roles:

1. Artifact Entity: $x : e_N \otimes_a \sigma$
 x exists because of event σ
2. Functional Entity: $x : e_N \otimes_t \tau$
the purpose of x is τ
3. Functional Artifactual Entity: $x : (e_N \otimes_a \sigma) \otimes_t \tau$
 x exists because of event σ for the purpose τ
 - a. *beer*: $(liquid \otimes_a brew) \otimes_t drink$
 - b. *knife*: $(phys \otimes_a make) \otimes_t cut$
 - c. *house*: $(phys \otimes_a build) \otimes_t live_in$

Artifactual Predicate Types

Predicates formed with Artifactual Entities as arguments:

1. *spoil*: $e_N \otimes_t \tau \rightarrow t$
2. *fix*: $e_N \otimes_t \tau \rightarrow (e_N \rightarrow t)$

- a. $\lambda x: e_A [\textit{spoil}(x)]$
- b. $\lambda y: e_A \lambda x: e_N [\textit{fix}(x,y)]$

- The beer spoiled.
- Mary fixed the watch.

Entities formed from the Naturals and Artifacts by a product type between the entities, i.e., the dot, •.

1. a. Mary doesn't believe the book.
b. John sold his book to Mary.
2. a. The exam started at noon.
b. The students could not understand the exam.

Motivating Dot Objects

When a single word or phrase has the ability to appear in selected contexts that are contradictory in type specification.

If a lexical expression, α , where $\sigma \sqcap \tau = \perp$:

1. $[\]_{\sigma} X$

2. $[\]_{\tau} Y$

are both well-formed predications, then α is a dot object (complex type).

Dot Object Inventory: 1

1. Act•Proposition: promise, allegation, lie

- I doubt John's promise of marriage.
- John's promise of marriage happened while we were in Prague.

2. Attribute•Value: temperature, weight, height, tension, strength

- The temperature is rising.
- The temperature is 23.

1. Event•Information: lecture, play, seminar, exam, quiz, test
 - a. My lecture lasted an hour.
 - b. Nobody understood my lecture.
2. Event•Music: sonata, symphony, song, performance, concert
 - a. Mary couldn't hear the concert.
 - b. The rain started during the concert.

Dot Object Inventory: 3

1. Event•Physical: lunch, breakfast, dinner, tea
 - a. My lunch lasted too long today.
 - b. I pack my lunch on Thursdays.
2. Information•Physical: book, cd, dvd, dictionary, diary, mail, email, mail, letter
 - a. Mary burned my book on Darwin.
 - b. Mary believes all of Chomsky's books.

Predicates formed with a **Complex Entity Type** as an argument:

1. *read*: $phys \bullet info \rightarrow (e_N \rightarrow t)$
2. Expressed as typed arguments in a λ -expression:
 $\lambda y: phys \bullet info \lambda x: e_N [read(x,y)]$
3. **Mary read the book.**

- Hearst 1992.
- Yamada and Baldwin 2004.
- Cimiano and Wenderoth 2005, 2007.
- Poesio and Almuhareb 2008.
- Baroni and Lenci 2008.

Lab on Qualia Extraction

- CQL regular expressions for the patterns discussed in class
- run queries on Ententen13 corpus through SkE
- extract relevant pairs

Formal-specific Constructions

- [tag=" N.*"][lemma=" such"][lemma=" as"][tag=" N.*"]
- [lemma=" such"][tag=" N.*"][lemma=" as"]
- [tag=" N.*"][lemma=" and"][lemma=" other"][tag=" N.*"]
- [tag=" N.*"][lemma=" or"][lemma=" other"][tag=" N.*"]
- [tag=" N.*"][word=" including"][tag=" N.*"]
- [tag=" N.*"][lemma=" especially"][tag=" N.*"]
- [word=" favorite"][tag=" N.*"][word=" is"][tag=" N.*"]

Constitutive-specific constructions

- `[tag=" N.*"][word=" is"][] {0,2}[word=" part"][lemma=" of"][] {0,2}[tag=" N.*"]`
- `[tag=" N.*"][word=" made"][lemma=" of"][tag=" N.*"]`
- `[tag=" N.*"][word=" consists"][lemma=" of"][] {0,1}[tag=" N.*"]`
- `[tag=" N.*"][word=" containing"][] {0,2}[tag=" N.*"]`

Telic-specific constructions

- [lemma="a"][tag="N.*"][word="worth"][tag="VVG"]
- [tag="N.*"][lemma="deserve" | lemma="merit"][tag="VVG"]
- [lemma="enjoy" | lemma="prefer"][tag="VVG"]
- [lemma="a"][tag="N.*"][word="for"][tag="VVG"]
- [lemma="a"][tag="N.*"][word="used"][word="for"]
[tag="VVG"]

- [tag="N.*"][tag="V.*D"][tag="V.*N|V.*D"]