

To Appear from Oxford University Press
2016

A Guide to
Generative Lexicon Theory

James Pustejovsky and Elisabetta Jezek

Chapter 2

Introducing Qualia Structure

In this chapter we introduce the mechanism used in GL to represent the core meaning of words. This system is based on four dimensions of meaning called *Qualia*. Qualia capture different properties of objects, as they are reflected in language: the *Formal* encodes taxonomic information, the *Constitutive* encodes information about the material and parts of objects, the *Telic* encodes information regarding the function and purpose, and the *Agentive* encodes information about the origin of an object. In the introduction of Qualia, we focus on nouns and touch only briefly on verbs in section 2.5. A thorough illustration of how Qualia can be used to represent the meaning of verbs and adjectives is developed later in Ch. 4. The goal of this chapter is to become acquainted with the Qualia formalism, to understand its motivating principles, and to be able to apply it to the analysis of novel words and expressions. Qualia Theory is introduced informally in this chapter and will be developed gradually throughout the book.

2.1 What is a Quale?

A Quale (singular of *Qualia*), from the Latin meaning “of what kind of thing”, is a term GL borrows from philosophy to indicate 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 (for example, for the noun *dog*, having fur, barking, tail wagging, licking, etc.), Qualia relations as defined in GL are those that are relevant for the way the word is used in the language. For example, our knowledge that bread is something that is brought about through *baking* is considered a Quale of the word *bread*;

this knowledge is exploited in our understanding of linguistic expressions, such as *fresh bread*, meaning “bread which has been baked recently.”

Qualia relations are also referred to as *qualia roles*. The word *role* recalls the notion of *semantic role* used in the domain of verbal semantics to indicate how the various entities associated with a predicate participate in the event expressed by that verb (agents, patients, experiencers and so forth). Qualia roles in GL were first conceived as an argument structure for nouns, and have since been extended to all the major categories.

Qualia encode aspects of a word’s meaning that are often attributed as world knowledge by contemporary linguistic theories, i.e., the knowledge we have about objects in the world due to human experience, as in the example of *bread* and *bake* above. In GL, the role of such knowledge is identified when it impacts the behavior of linguistic expressions in usage. We will clarify later how the distinction between lexical meaning and world knowledge is approached in the model.

2.1.1 Qualia and other Formalisms

The Qualia-based system developed in GL to represent lexical meaning is intended to overcome some of the main difficulties encountered by more traditional systems of meaning representation. Traditionally, word meanings have been described in terms of sets of features. The basic idea behind this view is that the meaning of a word is made up of smaller units, called features, components, or primitives (that is, elements that cannot be decomposed any further).

The general strategy to define the meaning of a word in this framework is called lexical decomposition. For example, a *table* may be defined in terms of features such as [inanimate], [concrete], [with legs] and so on. The character of the features that make up a word meaning may vary depending on the category of the word under examination: noun meanings appear to encode features such as [animate], [artifact], [countable], [portable], and [part-of(x)]; verb meanings are assumed to include abstract features such as [act], [cause], [result], [manner], [motion], and so on. Some linguists use a binary notation for these features, such as [+cause] and [-cause]. According to the analysis based on lexical decomposition, a restricted number of features, when appropriately combined, will suffice to define the meaning of all words belonging to a lexicon.

Decomposition into primitives has been very influential in lexical semantics, but it has various shortcomings. The shortcoming that concerns us here is related to the semantic flexibility shown by words; that is, their

ability to take on an indefinite variety of possible senses depending on the other words they combine with. This key aspect of word meaning is not easily dealt with in traditional decompositional frameworks. For example, does the verb *like* have two different meanings in “He likes my sister” and “He likes vanilla ice cream”; and if so, how is this difference to be represented in decompositional terms? No traditional decompositional system, no matter how elaborate or refined, is able to answer this simple question in a straightforward fashion.

2.1.2 Decomposition in GL

Although we may say that the GL view of lexical meaning is fundamentally decompositional, in the sense that it is based on the claim that words encode complex concepts that may be decomposed into simpler notions, the general method adopted in GL to define the meaning of words is inverted with respect to the traditional decompositional strategy discussed in 2.1.1. That is, instead of concentrating on how a word meaning may be decomposed, GL examines how a word meaning may or may not compose with other meanings, and how it changes in the different contexts. In other words, 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.

For example, with the noun *car*, different aspects of the object are highlighted in the contexts in (2): the car is seen a physical object in (a); as a vehicle in (b); as the part that actually runs and warms up (the car’s engine) in (c) and (d) ; as something that that can be locked (the door) in (e); as the parts that produce sound (the car’s wheels) in (f). Drawing on such linguistic evidence, GL assumes that all these aspects of *car* (physical object, vehicle, engine, door, wheel) must be considered part of the meaning of the entry *car* in the lexicon, and hence part of its lexical semantics.

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

The procedure adopted in GL to identify the meaning of words requires a system of lexical representation that allows words to change their meaning in different contexts, while maintaining the distinction between word

meaning and world knowledge: this is what qualia structure aims to accomplish.

2.1.3 Hidden Events

The first motivation for Qualia relations comes from the analysis of polysemous nominals and adjectives, as described in Pustejovsky and Anick (1988). Specifically, it was proposed that there is a *hidden event* (h-event) in the lexical representation associated with nouns denoting objects that are made for a particular purpose, such as *door* and *book*. A hidden event is defined as the characteristic activity that, when performed, realizes the purpose of the object. Some examples of hidden events for artifactual nouns are given below:

- (2) a. a *door* is for “walking through”
- b. a *window* is for “seeing through”
- c. a *book* is for “reading”
- d. a *beer* is for “drinking”
- e. a *cake* is for “eating”
- f. a *car* is for “driving”
- g. a *table* is for “putting things on”
- h. a *desk* is for “working on”
- i. a *pen* is for “writing with”

According to Pustejovsky and Anick, the reason for including a hidden event in the lexical representation of these nouns is that in certain syntactic contexts this event appears to be present in the interpretation, even though it is not expressed in the syntax. For example, in (4) what is “finished” is the activity of drinking, but this information is not overtly expressed:

- (3) They finished the beer. (drinking)

On the other hand, the hidden information is not arbitrary. Rather, it depends on the semantics of noun. For example, in (5) what is finished is the activity of eating, not of drinking.

- (4) They finished their cake. (eating)

A similar phenomenon can be observed in Adjective-Noun constructions in (6), where the adjective modifies an activity associated with the

noun. For example, a *comfortable chair* is a chair which is good to “sit in”, *comfortable shoes* are shoes that are good to “wear” or “walk in”, and so forth.

- (5) a. a comfortable chair (to sit on)
b. comfortable shoes (to wear, to walk in)

Finally, consider compounds like *dinner dress*, *dessert wine* and *dinner table* in (7). Also in this case, the interpretation entails one (or more) hidden events corresponding to the typical activities associated with the objects *dress* and *wine*: a *dinner dress* is a dress which is “worn” on a special evening occasion, a *dessert wine* is a wine which is “drunk” while “eating” a dessert, and so forth.

- (6) a. a dinner dress (wearing)
b. a dessert wine (drinking)
c. the dinner table (eating at)

As we will see, the hidden event introduced in Pustejovsky and Anick corresponds the Telic Quale developed in classic GL.

2.1.4 Qualia Structure

The notion that lexical items can store information relating to hidden events and activities associated with the word is a useful device for helping in the interpretation of linguistic expressions, as noted above. In Pustejovsky (1991), a more elaborated set of relations is proposed, in addition to the hidden event, to represent the meaning of nominals. These relations are called *Qualia*, and the system of relation defining a single concept is called *Qualia Structure*. Qualia Structure consists of four basic roles:

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

Each Qualia role can be seen as answering a specific question about the object it is associated with:

- *Formal*: What kind of thing is it, what is its nature?
- *Constitutive*: What is it made of, what are its constituents?
- *Telic*: What is it for, how does it function?
- *Agentive*: How did it come into being, what brought it about?

Taken together, the answers to these questions can help elucidate the meanings of words in the language. Adopting the typed feature structure representation introduced in Chapter 1, we can view the qualia structure of a lexical item, α , as the four features below in (8), where F=Formal, C=Constitutive, T=Telic and A=Agentive:

$$(7) \left[\begin{array}{l} \alpha \\ \text{QUALIA} = \left[\begin{array}{l} \text{F} = \text{what } \alpha \text{ is} \\ \text{C} = \text{what } \alpha \text{ is made of} \\ \text{T} = \text{function of } \alpha \\ \text{A} = \text{origin of } \alpha \end{array} \right] \end{array} \right]$$

Recalling the examples of *car* above, we can now identify which qualia role is exploited in the different contexts (9) and represent the results in the feature structures for F and C in (10).

- (8) a. This car weighs over 2,000 lbs. (car as material (C))
 b. We buy vehicles such as cars and buses. (car as vehicle (F))
 c. John started the car. (part of car, engine (C))
 d. You should warm your car up in winter. (part of car, engine (C))
 e. Did you lock the car? (part of car, door (C))
 f. The car screeched down the road. (part of car, wheel (C))

$$(9) \left[\begin{array}{l} \text{car} \\ \text{QUALIA} = \left[\begin{array}{l} \text{F} = \text{vehicle} \\ \text{C} = \{\text{engine, door, wheels, ...}\} \end{array} \right] \end{array} \right]$$

It should be pointed out that not all lexical items carry a value for each qualia role. Some are left unspecified, while others are populated with more than one value. For example, nouns denoting natural kinds (e.g., *rock, fish, air, sea*) typically do not have a value for the Agentive Quale, since

the objects they reference are not products of human creation. While such entities are obviously the product of natural forces, we assume an anthropocentric folk ontology. On the other hand, an artifact such as a *letter* is an entity brought about by a specific activity of “writing”, identified through the Agentive role:

$$(10) \left[\begin{array}{l} \textit{letter} \\ \text{QUALIA} = \left[\text{A} = \mathbf{write} \right] \end{array} \right]$$

The purpose of this same object can be identified through the Telic role, namely, identifying the purpose of the letter as “reading”.

$$(11) \left[\begin{array}{l} \textit{letter} \\ \text{QUALIA} = \left[\begin{array}{l} \text{T} = \mathbf{read} \\ \text{A} = \mathbf{write} \end{array} \right] \end{array} \right]$$

Similar remarks hold for other informational objects such as *book*, *novel*, and *diary*. In later discussions, we will see how to distinguish communicative artifacts such as *letter* from informational artifacts such as *book*.

Now consider a more elaborate example involving the lexical semantics associated with the noun *house*. First, observe the contexts where different Qualia of the noun are activated (13). Then, look at the proposed GL representation that follows from the observed data (14).

- (12) a. He owns a two-story house. (house as artifact (F))
 b. Lock your house when you leave. (part of house, door (C))
 c. We bought a comfortable house. (purpose of house (T))
 d. The house is finally finished. (origin of house (A))

$$(13) \left[\begin{array}{l} \textit{house} \\ \text{QUALIA} = \left[\begin{array}{l} \text{F} = \mathbf{building} \\ \text{C} = \{\mathbf{door, rooms, ...}\} \\ \text{T} = \mathbf{live.in} \\ \text{A} = \mathbf{build} \end{array} \right] \end{array} \right]$$

2.1.5 Criteria for identifying Qualia Values

In the discussion that follows, we will make a distinction between those qualia values that are lexically specified and those which are introduced by composition in the syntax. Lexically specified (viz., default) values are

identified by examining the distribution of the noun in context. For example, in the context of *shatter* in (15a) the noun *window* appears to be used to reference the part represented by the pane of glass, while in the context of *wooden* and *rotting* (15b) it refers to the frame. Because of this evidence, we might consider *pane* and *frame* as default values of the Constitutive quale for the noun *window*.

- (14) a. The rock shattered the window.
 b. Wooden windows are prone to rotting.

$$(15) \left[\begin{array}{l} \textit{window} \\ \text{QUALIA} = \left[\text{C} = \{ \textit{pane}, \textit{frame}, \dots \} \right] \end{array} \right]$$

Consider now the following contrast:

- (16) a. The museum is open until 6:00 pm.
 b. *The painting is open until 6:00 pm.

Notice that the noun *museum* in (17a) is being interpreted as the prototypical service event associated with the noun, namely an exhibition. This could be encoded in the Telic quale for *museum*, as shown below.

$$(17) \left[\begin{array}{l} \textit{museum} \\ \text{QUALIA} = \left[\begin{array}{l} \text{F} = \textit{institution} \\ \text{T} = \textit{exhibit} \end{array} \right] \end{array} \right]$$

However, while the noun *painting* might also be conventionally associated with exhibitions and museums, it does not permit of a similar interpretation in (17b). Such discriminative examples can be systematically used as linguistic evidence for determining what information is lexically associated with the qualia structure of a word. For example, “being exhibited in a museum” would not participate in the direct meaning of *painting* and hence does not constitute a qualia value of the word.

When language accesses the component parts of a word’s meaning with systematic regularity, there is reason to think that those parts might arguably be encoded in the lexical semantics for that word. For example, consider how the noun *car* frequently cooccurs with verbs denoting human actions in subject position: that is, a car can *travel*, *go uphill*, *honk*, *wait for somebody*, and so forth.

- (18) a. The car is waiting in the driveway.
 b. A car honked from behind.

Such examples are cases of *metonymy*, and are an interesting sense extension from the vehicle to an individual associated with it. Such evidence suggests that the information that a car has a driver is not only part of our world knowledge but is in fact encoded in the lexical entry and available for syntactic selection. In GL, this information is coded as an argument to the predicate that fills the Telic role of car, e.g., *drive*:

$$(19) \left[\begin{array}{l} \textit{car} \\ \text{QUALIA} = \left[\begin{array}{l} \text{F} = \textbf{vehicle} \\ \text{T} = \textbf{drive(human,vehicle)} \end{array} \right] \end{array} \right]$$

A related metonymic extension is seen with nouns such as *house* and *café*, which are often used to refer to the people who live in or work in the structure.

- (20) a. Do you want the whole house waken up?
 b. The rest of the house was sleeping.
 c. You had the whole café laughing.

Again, such data provide evidence for specific Telic values for these noun concepts; **live_in(human,building)** and **eat_in(human,building)**.

$$(21) \left[\begin{array}{l} \textit{house} \\ \text{QUALIA} = \left[\begin{array}{l} \text{F} = \textbf{building} \\ \text{T} = \textbf{live_in(human, building)} \end{array} \right] \end{array} \right]$$

2.1.6 Historical Note

The notion of Qualia in GL originates from the Aristotelian theory of explanation (*aitia*), usually known as the doctrine of four causes. *Aitia* is a Greek term (pl. *aitiai*) meaning ‘explanation’. One of the common ways to interpret the Aristotelian scheme is to see it in terms of causal links. According to this interpretation, an *aitia* is the cause of something (x is an *aitia* of y): for example, a sculptor is the *aitia* of a statue, a carpenter is an *aitia* of a table, and so on. In GL, however, Qualia Structure is derived from Moravcsik’s (1975) interpretation of the Aristotelian theory of *aitia*. Moravcsik proposes to interpret the theory of *aitia* as a theory of understanding instead of a theory of causation. According to this view, *aitiai* are not simply causes but rather modes of description of an object that allow us to understand it not only by knowing what it is (this is the information provided by the Formal Quale) but also by grasping how it functions (Telic), what its constituents

are (Constitutive), and what brings it about (Agentive). GL assumes that the four *aitiai* are relevant not only from an ontological point of view, as for Moravcsik, but also for the characterization of lexical meaning and the modes of compositionality in the grammar.

2.1.7 Qualia in Different Linguistic Phenomena

Explicit reference to qualia structure has proven to be quite useful for representing many linguistic phenomena related to polysemy and lexical ambiguity, which are difficult to deal with using traditional lexical representations. Some of these phenomena, discussed above, include:

- (22) a. Contextual modulations of noun meaning, due to the selecting predicate (2); *start/lock* the car
- b. Inference of implicit predicates from particular constructions:
 Verb-Noun (4, 5) (*finish* the beer / cake);
 Adjective-Noun (6) (*comfortable* chair / shoes);
 Noun-Noun (7) (*dinner* dress / table).

Additional phenomena include (a) the flexibility of light verbs support verb constructions), and (b) Noun-to-Verb transformations. The first can be seen with verbs such as *make*, *take*, and *have*, where specific information in the Qualia Structure of the complement is exploited in the overall interpretation of the construction, as illustrated in (24), where *take* is interpreted as *ingest* in the context of *tablet* and as *use to travel* in the context of *train*;

- (23) Light verbs specifications:
 i. Take a *tablet* (TELIC = ingest)
 ii. Take a *train* (TELIC = travel with)

Noun-to-Verb transformations, on the other hand, involve a category shift of a noun to a verb, in which the noun's Telic role becomes the verb meaning. Examples include the verbal forms of the nouns *fax*, *microwave*, and *lace*, as demonstrated below in (25).

- (24) Noun-to-Verb transformations:
 a. *fax* a document: (TELIC = send)
 b. *microwave* the chicken: (TELIC = cook)
 c. *lace* the shoes: (TELIC = tie)

There are many phenomena other than polysemy that receive a natural interpretation if Qualia are assumed as one of the basic components of lexical meaning. For example, the apparent necessity of adjuncts in constructions such as short passives (26), middles (27) and past participle constructions (28) may be analyzed as determined by the fact that the noun and the predicate in the construction form a *qualia pair*, that is, a combination in which the predicate expresses one of the qualia values of the noun (like *picture-paint*, *book-read*, or *house-build*), a phenomenon called *co-specification* in GL (cf. Chapter 5). In GL terms, qualia pairs are considered uninformative in a typical discourse context, since the verb predicates information which is already encoded as part of the noun's meaning. Typically, this is associated with properties of definiteness of the NP whose head is part of the qualia pair. Hence, this is why an adjunct is needed to avoid the sense of uninformativity in these uses, as illustrated in (26) below. Similar remarks hold for short middles (27) and adjectival uses of past participles in Adj-N pairs (28).

- (25) Short passives (TELIC(picture) = paint):
 a. *This picture was painted.
 b. This picture was painted *in 1604*.
- (26) Middles (TELIC(book) = read):
 a. *This book reads.
 b. This book reads *easily*.
- (27) Adjectival Use of Past Participles (TELIC(book) = read):
 a. *a built house;
 b. a *recently* built house.

Throughout the book, in fact, we will see how qualia structure is implicated in far more grammatical constructions and licensing operations than one would at first imagine.

2.2 The Four Qualia Roles

In the following sections, we discuss in more detail the properties of the individual Qualia introduced in the previous section. For now, we restrict our discussion to the semantics of nouns. For each role, we give a definition and provide examples that clarify how it is interpreted and how it can be used in lexical analysis. The illustration of Qualia Theory we present is

based on theoretical generalizations formed from the empirical analysis of lexical distributions in language.

2.2.1 Formal

A lexical item carries information about the basic conceptual category with which it is associated. In GL, this information is coded in the Formal quale (FORMAL) of the word. Particularly, the Formal quale establishes a relation between the entity denoted by a word (e.g., *dog*) and the category it belongs to (i.e., *animal*). It is this relation that enables one to grasp the nature of an entity by distinguishing it among other kinds. It answers the question: “What kind of entity is x?” For example, a *rock* is a natural kind, a *table* is an artifact, a *car* is a vehicle, a *park* is a location, *water* is a liquid, a *plant* is a living thing, a *fish* is an animal, a *hand* is body part, a *glass* is a container and so on. Sometimes, more classifications are possible for the same type of object: for example, a *knife* can be a weapon or a kitchenware. Moreover, classifications at different levels of generalization are available for reference: for example, in (29) *water* is seen as a liquid (its immediate superordinate), a fluid (a higher superordinate) or a substance (the highest superordinate). These are seen in the *such as*-construction, shown below.¹

- (28) a. a liquid such as water;
 b. fluids such as water or air;
 c. substances such as fluids, salts, glucose and carbon dioxide.

The hierarchy reflected by these constructions is illustrated in 2.1 below.

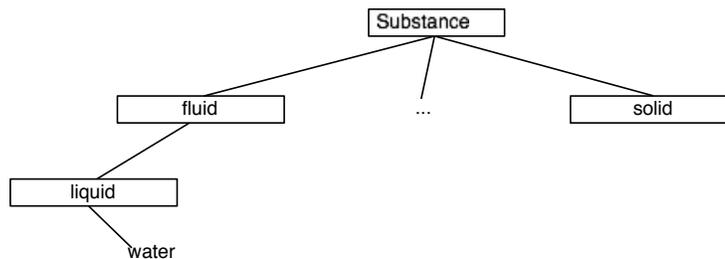


Figure 2.1: Lexical Hierarchy reflecting Type Structure

¹This construction has been used for lexical acquisition purposes in Computational Linguistics research (cf. Hearst (1992,1995), Lin (1998)), along with other grammatical patterns indicating specific function or semantic roles (cf. Pustejovsky et al, 1993, Wilks et al, 1996).

As another illustration, consider the representation of the Qualia structure of *rock* in (30), focusing on the Formal role, and the accompanying linguistic distributional evidence for such a classification in (31). Note the the examples in (31) give us information not only about the superordinate category (natural) but also about properties of rock: for example (31c) tells us that rocks are made up of hard material, a point we will come back below.

$$(29) \left[\begin{array}{l} \textit{rock} \\ \text{QUALIA} = \left[\text{F} = \mathbf{physical} \right] \end{array} \right]$$

- (30) a. inanimate objects such as rocks;
 b. natural elements such as rocks, soil, timber and so on;
 c. hard material such as rocks.

Categories associated with words may be analyzed as clusters of characteristic properties. Salient properties of a category (i.e., the properties which enter into the constitution of the concept for that category) answer the question, “What makes this y an x?”: for example, “What makes this animal *a dog*?”, “What makes this thing *a table*?” and so forth. We may identify salient properties for an entity by looking at the distribution of the noun that references it. For example, the example of *rock* in (31c) illustrates that rocks are conventionally classified as made of hard material, despite the fact that a distinction between hard rock and soft rock may be identified at a more technical level (marble is a hard rock, while chalk is a soft rock). This information, however, pertains to the Constitutive role, as it involves the material which rocks are made of. This will be discussed in the next section.

We will distinguish two kinds of properties that an object may have, for purposes of classifying this object relative to other entities:

- (31) a. Properties that are relevant for the physical taxonomic classification of an entity;
 b. Properties describing an entity’s internal and external constitution, typical behavior, use, purposes, or function.

The Formal role typically provides access only to the properties of an object that are relevant for its taxonomic classification. These are the properties that a word inherits from its superordinate. It is on the basis of these properties that we are able to draw class inferences. For example, by virtue of its Formal relation with liquid, water inherits the defining properties of

liquids, such as fluidity. Note that the constituents of water, namely hydrogen and oxygen, do not inherit the properties of water, since the relation they hold with water is not a Formal relation (is-a relation) but a Constitutive relation (made-of relation). For now, we will assume that property inheritance (class inference) only occurs along with Formal qualia values.

Some lexical items appear to have more than one Formal role value. These are cases of *multiple inherence* over the Formal quale: for example, the noun *book* refers to both a physical object (as in *hand me the book*) and an informational object (as in *believe the book*); the noun *lunch* refers both to an event (as in *during lunch*) and the food eaten (as in *a spicy lunch*); the noun *house* denotes both a physical object (as in *build a house*) and a location (as in *leave the house*). In the GL model, these are called *dot objects* and will be introduced in Chapter 3.

We can summarize the properties of the Formal Quale as providing access to three kinds of information:

- (32) a. The basic category associated with the word (i.e., its semantic type);
- b. The position of the word in the hierarchy of types following from this association;
- c. The salient properties which enter into the definition of the type, which are inherited by the word along the Formal role.

Pustejovsky (1991,1995) provide a list of Formal attributes (or *factors*) for the class of nouns denoting concrete entities. These include physical characteristics such as the following:

- (33) a. Spatial characteristics, intrinsic orientation;
- b. Size and dimensional properties;
- c. Shape and form;
- d. Color.
- e. Position.

Each attribute may be filled by a value: for example, in *a red car*, *red* is the value (filler, descriptor) of the Formal attribute “color”; in *a long dress*, *long* is the value of the Formal attribute “dimension”; in *a round table*, *round* is the value of the Formal attribute “shape”, and so forth. Some combinations may be ambiguous: for example, a *red pen* might mean “a pen that is red” or “a pen that writes in red ink”. In the first case, we assume that the adjective activates the Formal attribute “color” of the noun, while in the

second case, it appears that it is the Constitutive or the Telic which play a role in the interpretation. Obviously, several modifications, referring to different factors, might be present in the structure of a single NP, such as in (35) below, where FACTOR_F refers to a named factor of the Formal role:

(34) a *large* (Size_F) *round* (Shape_F) table

Conversely, a named factor value can generally be filled only once for each modified noun.

(35) *a *round* (Shape_F) and *square* (Shape_F) table

Lexical meaning often provides default values for the different Formal factors or attributes. Default values are inherent properties of entities which enter into their denotation. For example, we might assume that the “Size” value associated with the noun *ant* is *small*, when evaluated relative to the superordinate class for the noun *insect*. This can be supported from such uses as seen in (37):

(36) ... an *insect* as small as an *ant* ...

Default values, however, may be updated from discourse context in composition: for example, in *large ant*, context makes us update the value of the Size factor from *small* (default) to *large* (*for an ant*). Comparison classes of adjectives are suggested by specific information from the Formal role; e.g., in *a large ant* vs. *a small dog*, we have basic category and ontological classification information specified that gives us a way to constrain the interpretation of relative interpretations of “Size” in such constructions.

It is often the case that there is no value specified lexically for a particular Formal factor. For example, one might assume that part of our knowledge of ants is that they are typically black in color. Modeling such information lexically entails specifying a default value for the color factor associated with the Formal role associated with the lexical item *ant*. For many natural kinds, however, no such default color can be associated with the individuals associated with that kind, e.g., dogs, cats, rocks. In such cases, the value for a particular Formal factor may be introduced in context. For example, in *black dog*, the expression compositionally introduces the value *black* for the color factor associated with the Formal for the noun *dog*.

As mentioned above, the Formal role can be seen as a containment relation: if the Formal(water) = liquid, then the relational form can be read as *is-a*(water, liquid). The two terms participating in the Formal relation

tend not to cooccur in the same local context since they are in a paradigmatic relation with each other. We may, however, occasionally find both terms in constructions such as those shown in (38).

- (37) FORMAL-specific Constructions:
- a. NP *such as* NP: *events such as lectures, walks, tours and meetings*;
 - b. *such* NP *as* NP: *such areas as children's playground*;
 - c. NP *and other* NP: *rum and other spirits*;
 - d. NP *or other* NP: *insects or other animals*
 - e. NP, *including* NP: *recyclable materials including glass*;
 - f. NP, *especially* NP: *cool temperate countries especially Europe and North America*;
 - g. *favorite* NP *is* NP: *Mario's favorite food is pasta*.

On the other hand, one regularly finds Formal factors expressed through adjectival, nominal, verbal or prepositional descriptors, as illustrated in (39-42):

- (38) a. a *flat* screen (Shape_F)
 b. a *leaning* tower (Shape_F)
 c. a *straight* line (Shape_F)
 d. a *thick* sweater (Dimension_F)
 e. an *expensive* car (Cost_F).
- (39) a. the *height* of the wall (Dimension_F)
 b. the *length* of the table (Dimension_F)
 c. the *depth* of the sea (Dimension_F).
- (40) a. the *front* of the house (Orientation_F)
 b. the *top* of the table (Orientation_F)
 c. the *foot* of the stairs (Orientation_F)
 d. the *head* of the bed (Orientation_F)
 e. the *facade* of the building (Orientation_F);
- (41) a. *clean* the table (Surface_F)
 b. *wipe* the floor (Surface_F)
 c. *paint* the wall (Surface_F).

In this section, we have focused mostly on the basic properties of the Formal role for lexical items denoting natural kinds. In later chapters, we will examine how the Formal is defined for words denoting more complex

entities, including objects with specific functions and purposes (artifacts), objects which inherit from multiple superordinates (dot objects), as well as relational nouns.

2.2.2 Constitutive

In the previous section, we saw how the representation of word meaning encodes the basic categorical knowledge associated with a lexical item through the Formal role of the qualia structure. The classification resulting from this information is essential for differentiating objects within a class. To make this clear, consider how we analyzed the noun *water*. We said *water* was defined in terms of its relation to the supertype, *liquid*, which in turn was defined as a *fluid*, and so on, through the Formal role. Similarly, the noun *woman* has a Formal value of *human*, which in turn has a Formal value of *animal*, and so on.

Upon closer examination, however, it becomes clear that the values associated with the Formal roles for these two cases are quite different: namely, 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 in fact known as the “count/mass” distinction in linguistics, and is a key aspect in determining the grammatical behavior and distribution of nouns in language. Because this distinction is so central to the conceptualization of our world, it plays a central role in how the Formal role for nouns is interpreted in a lexicon.

In GL terms, there is not actually a mass/count distinction in the grammar. Rather, this distinction is an artifact of how the Formal role interacts with the *Constitutive* role for a lexical item. To understand how, we look at a range of examples somewhat informally. We can think of the Constitutive (CONST) role associated with an object as encoding information about what is “inside” that object, the material the object is made of (i.e., its stuff), and the parts it consists of.

For example, the noun *car* denotes a vehicle (its Formal role value), but is made of many different parts, e.g., chassis, engine, seats, and so on. We say that these parts denote the value of the Constitutive of the noun *car*. Similarly, a tree has many parts, such as a trunk, branches, leaves, and roots. So, while the noun *tree* denotes a plant (its Formal role value), these aspects would designate the value of the Constitutive role for *tree*.

Let us now return to mass terms such as *water* and *air*. One of the defining characteristics of a mass term is that smaller parts of the material being referred to are still that material. In other words, any decomposition

results in the same value. This has a very elegant representation in GL, where count and mass terms are distinguished by their qualia structures. We will make the following distinction here. There is formally one relation of *part_of*(x, y), but when the types of x and y are identical, $x : \tau = y : \tau$, then we essentially have the *made_of* relation. When the types between x and y are distinct, however, $x : \tau = y : \sigma, \tau \neq \sigma$, then we have a *part_of* relation between heterogeneous objects. This is shown in (43) and (44) below.²

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

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

(43) MASS NOUN:

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

Count nouns have parts and possibly made of relation, and mass terms have no distinction between made of and parts.

The constraint in (44) is known as the *Formal-Constitutive Equivalence Constraint*. Another representation for the qualia structure for mass terms that we will use is: $[\text{F/C} = \alpha]$.

Now let us revisit the qualia structure for the noun *water* from the previous section. As a mass noun, there is no distinction made between the Formal and Const qualia, and we will represent it as shown in (45).

$$(44) \left[\begin{array}{l} \textit{water} \\ \text{QUALIA} = \left[\text{F/C} = \mathbf{liquid} \right] \end{array} \right]$$

Notice that this ensures that any type-based inference through the Formal role is still available to words denoting concepts pertaining to “stuff”: i.e., water is a liquid, liquid is a fluid, and so on.

The role played by the Constitutive quale would be significant enough from a semantic point of view if this were its only contribution to linguistic modeling. This is, however, only a small part of the descriptive power of this role. Since most nouns in the lexicon are count nouns (at least in English), the Formal-Constitutive Equivalence Constraint will not be relevant.

²In later chapters, we will use a formal device called *reentrancy*, which allows multiple feature labels (e.g., qualia roles) to point to the same value, e.g., α .

This means that the Constitutive is available for encoding concepts associated with a broad range of *part-of* relations. That is, if “*x* has a Constitutive value of *y*, then *part-of*(*y*, *x*).”

For example, for the noun *car*, the Constitutive expresses the relation of the car to its constituent elements, e.g., chassis, engine, windows, seats, and so on.

$$(45) \left[\begin{array}{l} \textit{car} \\ \text{QUALIA} = \left[\begin{array}{l} \text{F} = \mathbf{vehicle} \\ \text{C} = \{\mathbf{chassis, engine, seats, windows, ...}\} \end{array} \right] \end{array} \right]$$

Besides expressing the internal makeup of an object, the Constitutive may also introduce the relation between an object and the larger object it is logically a part of, if such a dependency exists³. The noun *roof*, for example, denotes part of a larger entity, namely the concept of a *building*. This is characterized by reversing the direction of the Constitutive to allow reference to entities which are themselves parts of things. We will call this relation the *Inverse Constitutive* (C_I). Hence, if “*x* has an Inverse Constitutive value of *y*, then *part-of*(*x*, *y*).”

$$(46) \left[\begin{array}{l} \textit{roof} \\ \text{QUALIA} = \left[\begin{array}{l} \text{F} = \mathbf{phys} \\ \text{C}_I = \mathbf{building} \end{array} \right] \end{array} \right]$$

Notice that both of these aspects of the Constitutive role come together in the semantics of a noun such as *room*, which is both part of something (a building) and has parts in it (walls, floor, ceiling). This can be expressed in the lexical representation shown below in (48).

$$(47) \left[\begin{array}{l} \textit{room} \\ \text{QUALIA} = \left[\begin{array}{l} \text{F} = \mathbf{space} \\ \left[\begin{array}{l} \text{C} = \{\mathbf{walls, floor, ceiling, ...}\} \\ \text{C}_I = \mathbf{building} \end{array} \right] \end{array} \right] \end{array} \right]$$

The linguistic relevance of the Constitutive quale for the noun *room* can be seen in (49), where in the context of *paint*, *room* is able to assume the interpretation of “the walls from the room”, while in the context of *sweep* it refers to the “floor of the room.”

- (48) a. John was going to paint his room ([CONST = walls]).
 b. She has swept the room ([CONST = floor]).

³This is explored in Pustejovsky and Anick (1988) and Vikner and Hansen (1994).

Similarly, in the selectional context of the verbs *cut* and *sharp*, the noun *knife* can be used to denote specifically the part that enables the cutting, i.e., the blade (50).

- (49) a. The knife cut his finger ([CONST = blade]).
 b. a sharp knife ([CONST = blade])

$$(50) \left[\begin{array}{l} \textit{knife} \\ \text{QUALIA} = \left[\begin{array}{l} \text{F} = \textbf{phys} \\ \text{C} = \{\textbf{blade, handle}\} \end{array} \right] \end{array} \right]$$

Defining what counts as a part of an entity is a controversial topic. The matter is hotly debated especially within the theory of mereology (the study of parthood relations). In a way, the word *part* may be used to indicate any portion of a given entity, regardless of whether, for example, that portion is attached to the rest of the object, as with “the *handle* of a door”, or undetached, as “the *cap* of a pen.” However, among logicians, philosophers and ontologists it is often assumed that the “legitimate” parts of an object are those that demonstrate the following characteristics:

- (51) a. Parts are available in discourse as individual units;
 b. Parts make a functional contribution to the entity;
 c. Parts are cognitively salient.

The last point is meant to rule out arbitrarily demarcated regions or portions of an object, such as *the lower part of the wall*, which is clearly part of the wall, but not to be identified as a legitimate “part” of the wall, since it refers to a part of the wall with no clear boundaries. It is also generally assumed that a well-formed part-of relation should consist of elements of the same general type; a part of some physical entity will also be a physical entity, part of some time period will also be a time period, and so forth. According to this view then, the “weight” of a body does not figure among its parts, even though it will constitute an essential attribute (or factor) of that body.

The GL approach to the parthood relation expressed by the Constitutive and Inverse Constitutive qualia is driven by the grammatical reflexes that the constituency of an entity expresses, in the way we reference that entity in language. That is, in GL it is assumed that the Constitutive quale specifies only those parts of an entity that are relevant for the linguistic behavior of the noun expressing the whole, as the *walls* of a room. On this

view, entities may have several parts, only some of which may be recognized as values of the Constitutive quale for the word denoting that entity. Other parts may not be lexically encoded as part of the qualia structure of the word, even though they are semantically implied as parts of the entity. Therefore, it will be useful to distinguish between the following two noun types:

- (52) a. Nouns that specify a value for CONST lexically;
 b. Nouns that lexically lack a value for CONST.

Typically, nouns of type (53a) are words that lexical whole objects that allow for reference to parts, such as *room* or *knife*. These nouns may express one of the default values of CONST syntactically, as for example in a genitive construction (54a), or in a construction where the first term introduces reference to a part of the entity denoted by the second term (54b):

- (53) a. John was going to paint the room's walls.
 b. John was going to paint the walls of the room.

The behavior observed above follows from the fact that the noun *wall* is lexically specified within the qualia structure of the noun *room*. Notice however, that there are many ways to dynamically elaborate parts of objects that are not lexically specified in the noun. For example, although most rooms would arguably have corners as spatially defined partitions within the region defined by the room, there is no justification for calling a corner a "part of" the room, assuming the methodology we have adopted here. Consider the sentence in (55), where a part of the room has been identified compositionally in context:

- (54) He was standing in the corner of the room.

Although the most conventional choice for the Constitutive relation is the part-of relation, it can also be used in a broader sense to designate the relation of material constitution (made-of relation). This interpretation dates back to the Aristotelian scheme, where, for example, marble is described as being in a constitutive relation to the statue it is part of. In this sense, nouns such as *river*, *lake*, and *sea* have CONST values of water, since they are largely composed of this material. The qualia structure for the noun *river* encodes this value in its CONST role directly, as illustrated below, where we assume the FORMAL value for *river* is region or space, filled with the material designated by the CONST value.

$$(55) \left[\begin{array}{l} \textit{river} \\ \text{QUALIA} = \left[\begin{array}{l} \text{F} = \text{space} \\ \text{C} = \text{water} \end{array} \right] \end{array} \right]$$

While the FORMAL role can be modified by spatial predicates (cf. (57)),

- (56) a. They crossed the river. ([FORMAL = space])
 b. The river is wide. ([FORMAL = space])

in fact, the constitutive value is directly referenced by the verbs *freeze* and *pollute* as shown below.

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

While the example in (58b) exploits the qualia information associated with the CONST value of the noun *river* in order to arrive at the interpretation of “water being polluted”, CONST values of a noun can be explicitly expressed in syntax, as demonstrated in (59).

- (58) a. The water in the river is polluted.
 b. polluted river water.

Another typical way of expressing the made-of relation in English as well as in other languages is through nominal compounding. In this construction, the value for CONST is generally introduced in composition as a modifying bare noun. For example, in the Noun-Noun compounds in (60), the first noun (*plastic, paper, leather*) expresses the value of the CONST quale for the head of the expression: e.g., a plastic bag is a bag made of plastic, a paper cup is a cup made of paper and so on.

- (59) a. plastic bag
 b. paper cup
 c. leather shoes
 d. gold watch
 e. milk chocolate

$$(60) \left[\begin{array}{l} \textit{plastic bag} \\ \text{QUALIA} = \left[\begin{array}{l} \text{F} = \text{bag} \\ \text{C} = \text{plastic} \end{array} \right] \end{array} \right]$$

The made-of relation may also be introduced in adjectival modification constructions, such as in the Adjective-Noun constructions in (62):

- (61) a. a golden ring;
 b. a wooden floor;
 c. a metallic paint.

Finally, the made-of relation may be introduced indirectly, i.e., by referencing specific attributes related to the material. For example, with the expression *heavy chain*, we refer to the property "weight" that inheres of the object by virtue of its constitutive material; a chain may be *light* or *heavy* depending on the material it is made of. Therefore, we say that the word *heavy* introduces the value "heavy" to the attribute (or factor) "weight" of the CONST in the Qualia Structure of *heavy chain*. This is represented below, where the adjective *heavy* introduces the material of which it is measuring the weight.

$$(62) \left[\begin{array}{l} \text{heavy chain} \\ \text{QUALIA} = \left[\begin{array}{l} \text{F} = \mathbf{chain} \\ \text{C} = \mathbf{heavy(material)} \end{array} \right] \end{array} \right]$$

Similar remarks hold for other substance modifying attributes, such as "temperature". Consider, for example, the expression *cold water*, where *cold* modifies the factor "temperature". Because *water* is a mass term, the FORMAL value is identical to its CONST value. For a mass noun such as *soup* in the expression *thick soup*, the adjective *thick* is modifying a property of the density or viscosity of the liquid.

To conclude this section, we review some of the constructions seen in language that are specific indicators of constitutive relationships. These are shown in (64), where $Const(x, y)$ indicates the relation "*x* is the CONST of *y*".

- (63) CONST-specific Constructions:
 a. N_1 's N_2 : $Const(N_2, N_1)$
the room's wall;
 b. N_2 of N_1 : $Const(N_2, N_1)$
the door of the car;
 c. NP_2 is a part of NP_1 : $Const(NP_2, NP_1)$
brain is a very sensitive part of the body;
 d. NP_1 made of NP_2 : $Const(NP_2, NP_1)$
monuments made of stone and marble

- e. NP₁ of NP₂:
house of wood; $Const(N_2, N_1)$
- f. NP₁ consists of NP₂: $Const(NP_2, NP_1)$
the orchestra consists of ninety performers;
- g. NP₁ containing NP₂: $Const(NP_2, NP_1)$
a forest containing dead trees.

2.2.3 Telic

In this section, we turn to the analysis of the Telic Quale. The Telic Quale, TELIC (from the Greek term $\tau\epsilon\lambda\omicron\varsigma$, meaning “end” or “goal”), encodes information about the intended use or function of an object. Specifically, it expresses the relation that allows us to grasp what an entity is by knowing what it is used for, and what we normally do with it. Therefore, the value expressed by the Telic relation of a noun often corresponds to the activity in which the object named by the noun is typically involved; for example, a food item like *cake* has the Telic value of *eat*, while an instrument for writing such as *pen* has a Telic value of *write_with*.

$$(64) \left[\begin{array}{l} \textit{cake} \\ \text{QUALIA} = \left[\begin{array}{l} \text{F} = \textbf{food} \\ \text{T} = \textbf{eat} \end{array} \right] \end{array} \right]$$

$$(65) \left[\begin{array}{l} \textit{pen} \\ \text{QUALIA} = \left[\begin{array}{l} \text{F} = \textbf{tool} \\ \text{T} = \textbf{write_with} \end{array} \right] \end{array} \right]$$

The activity specified in the Telic should be interpreted as a potential activity or characteristic property of the object; as such, it is modally subordinated to the specific context which enables the possibility of this function to be activated. For example, a *pen*, has the characteristic of being used for writing, even if this characteristic is not exploited in a specific context. In similar fashion, *mail* is for sending even if it is not sent at all. This characteristic is a persistent property of the object, i.e., a property that persists through time. What is important for the Telic relation, however, is that when the activity is performed, the main purpose of the object is satisfied. We will examine in more detail this aspect of the Telic quale in the section dedicated to the relation between Telic and Modality in Chapter 5.

There is a good deal of empirical evidence suggesting that certain nouns encode information about the intended use or function of the object to which they refer (see also the examples in section 2.1.3). For example the

contexts in (67) clearly evoke a writing event made available by the noun *pen* even if this event is not expressed in the syntax. That is, the contexts in (67) are meaningful only if a writing event is assumed in the semantic representation of the sentence.

- (66) a. This pen does not work well. (does not write)
 b. Can I use your pen? (for writing)
 c. Have you got a red pen? (ambiguous, which writes in red)

Similarly, an eating event is made available by the noun *cake* in all the contexts in (68).

- (67) a. Any chocolate? Not after that cake! (after eating)
 b. I prefer cake to biscuits. (prefer eating)
 c. We skipped the cake and settled for another coffee. (skipped eating)

At times, the implicit predicate does not satisfy the intended function of the object but expresses one of the activities associated with its purpose. For example in the contexts in (69), the implicit predicate appears to be the “departure” of the train, while the relevant Telic value associated with a train would be that of *travel*.

- (68) a. There’s no train till 7:00 pm. (there is no departing)
 b. The train was delayed for an hour. (the departure was delayed)
 c. I left in time to catch the early train. (departing early)

In GL, the examples in (69) do exploit information relating to the function of the noun *train*, but it is more indirect than in our previous examples. In these cases, the elided predicate seems to relate to the *departure* of the train rather than simply *traveling* on it. Such knowledge is not directly part of the qualia structure per se, but is derived from it, by means of *projective* operations over these values. In other words, the act of “departing” is a projective activity, and is available for interpretation as one of the components of the overall Telic activity of “transportation”. These projected activities are treated as conventionalized attributes in GL. The representation of conventionalized attributes is discussed in detail in section 2.3, while a discussion of the mechanism of projection is presented in Chapter 5.

Another important piece of evidence for the Telic relation comes from Adjective-Noun constructions where the meaning of the adjective provides information regarding the intended use and function of the noun’s referent.

For example, in the context of *customer* in (70) *next* is referencing an ordering over the servicing of individual patrons (the customers), i.e., 'next to be taken care of', while in the context of *slide* it references the order of the viewing or presentation of the individual slide, i.e., 'next to be projected'.

- (69) a. the next customer (to be taken care of)
c. the next slide (to be projected)

Within GL, it is assumed that the contextual variation in meaning of *next* in (70) is determined by the fact that the adjective targets a sub-portion of the semantics of the head noun, particularly its Telic relation. The specific mechanism allowing this modification will be detailed in Chapter 5. Notice that context can be an important determining factor with such constructions. Imagine a conversation in a moving vehicle, where the passenger tells the driver to "turn left at the next intersection." In this case, *next* refers to the ordering imposed by the path assumed by the moving car.

Further examples in which the Telic relation of a noun plays a role in the interpretation of the modifying adjective (in both attributive and predicative position) are given below:

- (70) a. This is a difficult problem (to solve).
b. This is a difficult question (to answer).

The Telic of *problem* can be said to involve its solution, while the Telic of *question* involves its answer. The adjective *difficult* modifies the Telic aspect of the noun. Similar behavior is seen with adjectives such as *ready* in (72).

- (71) a. Your coffee is ready (to drink).
b. There's some lunch ready in the kitchen (to eat).
c. The car is ready (to drive).

In (73) we give a list of adjectives selecting the information made available by the Telic quale of the noun. For each adjective, we give a contextualized example, and indicate in brackets the telic value which is being modified.

- (72) Telic selectors:
fast food (to eat), a slow oven (to cook), a short novel (to read), a complex question (to answer), an easy place (to get to), useful, an effective antibiotic (to cure), agreeable, avoidable costs (to pay), enjoyable, a good doctor (to heal), a bad singer (to listen to), an interesting book (to read), ready meals (to eat).

Implicit predicates in Adjective-Noun constructions may not match the Telic information of the noun while at the same time are clearly associated with it. For example in (74) the adjective *heavy* appears to modify the activity of *digesting* instead of the activity of *eating*. This is shown by the fact that we can paraphrase *heavy* with ‘hard to digest’ rather than ‘hard to eat’.

(73) Heavy foods such as dairy products and meat.

In GL terms, this is again a case of exploitation of one of the *conventionalized attributes* associated with the Telic, which we will examine in more detail shortly.

Finally, as pointed out in 2.1.3, the Telic relation appears to play a role in the interpretation of nominal compounds. For example, in the compounds in (76) the Telic information encoded in the head nouns (e.g., *shop*, *glass*, *bus*, *train*) makes available the semantic relation between head and non-head that is exploited in the interpretation of the compound: a *book shop* is a shop where book are “sold”, a *wine glass* is a wine for “holding” wine, a *school bus* is a bus for “transporting” kids to school, and so forth.

- (74) a. a book shop (selling)
 b. a wine glass (holding)
 c. a school bus (transporting)
 d. a freight train (transporting)

In many constructions, the Telic information encoded in the non-head noun interacts with that of the head and constrains the interpretation. Compare *tooth brush* (cleaning) vs. *hair brush* (combing) in (75).

- (75) a. a tooth brush (cleaning)
 b. a hair brush (combing)

The information expressed by the Telic is characteristic of artifactual objects, i.e., objects created for a particular purpose. However, this information appears to be present also in nouns denoting a variety of concepts:

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

Functional locations, for example, are locations that are designed for the performance of particular activities. These activities may be encoded as Telic information of the corresponding nouns and may be activated in particular contexts. For example, in the context of *attend* in (78), *church* references the activity specified in its Telic quale and is reinterpreted as “church service” contextually, i.e., the event of the mass.

(77) He no longer attends church. (mass)

The Telic value of a noun denoting an artifact may be populated with more than one item or it may have none. In other words, not all nouns denoting artifacts encode Telic information. According to Asher and Pustejovsky (2006), this is because not all functions that we ascribe to artifacts “make it over to the lexicon”. That is, while all artifacts by definition have a characteristic function corresponding to the purpose for which they are built, information about this function is not necessarily incorporated in the meaning of the corresponding lexical item. We will come back on this issue, which touches on the relationship between the conventional aspects of word meaning and the general knowledge we associate with objects, in Chapter 3.3.

There are two main types of Telic, as discussed in Pustejovsky (1995): Direct Telic and Purpose (or Indirect) Telic. With the Direct Telic, the entity denoted by the noun is realized as the object of the activity (e.g., *cake* is the object of the Telic value, *eat*), while with the Purpose Telic, it is realized as subject or indirect object (e.g., *pen* is the subject or indirect object of the Telic value *write*). The Direct Telic characterizes the entity as something which one acts on directly. The Indirect Telic characterizes the entity as something which is used for facilitating a particular activity or has the function of carrying it out. Within the Indirect Telic, Busa et al. (2001) distinguish two subtypes: (i) the entity is an instrument (e.g., *pen* relative to *write*); (ii) the entity is an agentive noun (e.g., *singer* relative to *sing*). This information may be encoded in the lexical representation of *cake*, *pen* and *singer* as follows:

$$(78) \left[\begin{array}{l} \textit{cake} \\ \text{QUALIA} = \left[\begin{array}{l} \text{F} = \mathbf{food} \\ \text{T} = \mathbf{eat(human,food)} \end{array} \right] \end{array} \right]$$

$$(79) \left[\begin{array}{l} \textit{pen} \\ \text{QUALIA} = \left[\begin{array}{l} \text{F} = \mathbf{tool} \\ \text{T} = \mathbf{write_with} \end{array} \right] \end{array} \right]$$

$$(80) \left[\begin{array}{l} \textit{singer} \\ \text{QUALIA} = \left[\begin{array}{l} \text{F} = \mathbf{human} \\ \text{T} = \mathbf{sing(human, song)} \end{array} \right] \end{array} \right]$$

There are a number of Qualia-specific constructions for the Telic, that is, constructions where the two terms of a Telic relation co-occur. When the Telic activity being expressed corresponds to the Telic value specified in the noun (i.e., to the default Telic), the resulting expression is what is called a qualia pair. A qualia pair may take the form of a verb-noun pairing (82a), an Adjective-Noun pairing (82b), or a compound (82c):

- (81) a. The child drank some water.
 b. drinkable water;
 c. drinking water.

Notice that *-able* adjectives appear to impose a specific interpretation on the Telic activity of the noun they modify. For example *drinkable* does not simply mean ‘that can be drunk’ but rather ‘that is good for drinking’. Other examples of constructions where *-able* adjectives expressing the Telic value of the head noun are given below:

- (82) a. The rent is payable monthly.
 b. a very readable text-book;
 c. a very playable game.

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. Consider, for example, the complex nominal *shopping bag*. The lexically specified Telic value for *bag* is *hold*, by virtue of it being a container; this value is updated in composition by the modifier *shopping*, which becomes the Telic value of the overall expression.

$$(83) \left[\begin{array}{l} \textit{shopping bag} \\ \text{QUALIA} = \left[\begin{array}{l} \text{F} = \mathbf{container} \\ \text{T} = \mathbf{shopping} \end{array} \right] \end{array} \right]$$

Similar remarks hold for the compounds listed below.

- (84) a. ironing board (used for ironing)
 b. swimming pool (used for swimming)

- c. dining room (used for dining)
- d. frying pan (used for frying)
- e. cutting knife (used for cutting)

Other specific constructions expressing Telic values (either lexically specified or updated compositionally) are given below in (86), where *Telic*(*x*, *y*) indicates the relation “*x* is the TELIC of *y*”.

- (85) TELIC-specific Constructions:
- a. an NP to V: *Telic*(*V*, *NP*)
a book to read;
 - b. an NP worth V-ing: *Telic*(*V*, *NP*)
a question worth asking;
 - c. the NP merits/deserves V-ing: *Telic*(*V*, *NP*)
This book deserves reading.;
 - d. enjoy/prefer V-ing NP: *Telic*(*V*, *NP*)
enjoy listening to music / prefer watching television;
 - e. an Adj NP to V: *Telic*(*V*, *NP*)
a difficult question to ask;
 - f. an NP (used) for V-ing: *Telic*(*V*, *NP*)
a spade (used) for digging;

Natural Telic

While the Telic of an artifact gives information about the intentional activities that satisfy the object’s design or purpose (*pens* are for *writing*), the Telic of a natural kind (*human*, *dog*, *water*, and so forth) encodes information about the actions and properties that the object engages in, but that are not in any way intentional or purposive. For example, when we describe the analytic properties associated with humans, as expressed in (87), we are predicating a “natural Telic” property of that entity.

- (86) a. Humans breathe/think.
b. Rivers flow.
c. The heart pumps blood.

It is not the intentional purpose of a heart to pump blood, but it is a necessary activity for the object so defined. Likewise, 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.

The linguistic relevance of the property of “flowing” for the noun *river* appears in (88), where in the context of *swift, lazy* etc., *river* allows for reference to its Natural Telic value: in other words, a *swift river* is a river ‘which flows at high speed’. Additional examples modifying the attribute of “speed of flow” are shown below.

(87) a fast/rapid/slow/lazy river (flowing)

Similarly, corpus analysis shows that the expression *flowing river* is uncommon while expressions specifying the manner of flowing are totally normal and frequent, as seen in (89).

(88) a. a fast/quietly/slowly/steadily flowing river;
b. a gently/peacefully flowing river.

This evidence supports the GL view that *flowing* is part of the meaning of *river*, and hence is uninformative as a modifier to the noun *river*. By contrast, in the expressions in (89) the modifier introduces new conceptual material, thus satisfying the requirement that every linguistic expression must be informative in the discourse context (see 2.1.7).

Given such behavior, we now return to the qualia structure for the noun *river*, and supplement the Formal and Constitutive values with a specification for the Natural Telic, which we designate as T_N , namely the value *flow*. This is illustrated in (93) below.

(89)
$$\left[\begin{array}{l} \textit{river} \\ \text{QUALIA} = \left[\begin{array}{l} \text{F} = \text{space} \\ \text{C} = \text{water} \\ \text{T}_N = \text{flow} \end{array} \right] \end{array} \right]$$

The naturally occurring activities described above are to be distinguished from those associated with intentionally created or designed objects, such as letters, pens, knives, etc.

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), as in the case of *letter* in 2.1.4. Natural kinds lack this association, as they do not encode an Agentive value.

There are, of course, purposes and functions that can be attributed to natural kinds, most notably the recognition that something can be consumed as food, or ingested as a beverage. Hence, although we classify apples as fruits, they are, more importantly, *edible fruits*. Hence, just as we

can “recover” the implied Telic for an artifactual noun, such as *cake* in (92a), we can perform the same computation in (92b), given our classification of apples as foods.

$$(90) \left[\begin{array}{l} \textit{apple} \\ \text{QUALIA} = \left[\begin{array}{l} \text{F} = \mathbf{fruit} \\ \text{T}_N = \mathbf{eat} \end{array} \right] \end{array} \right]$$

- (91) a. Mary enjoyed the cake. (TELIC = eat)
 b. Mary enjoyed the apple. (TELIC = eat)

In cases such as (92b), there is no Agentive-Telic pairing but there is an intention associated with the activity that has been given to the natural entity, namely, eating for sustenance and drinking for quenching, respectively, in the above examples. We will examine these cases in more detail in Chapter 5.

2.2.4 Agentive

In this section we examine in more detail the Agentive Quale, *A*, a role that encodes information about the origin of an object, or its “mode of coming into being”, to use Aristotle’s terminology. This is a crucial role for differentiating the kinds of objects, properties, and relations that exist in the world, since it provides a mechanism for distinguishing natural entities from non-natural entities. As cognitive agents interacting with the world, we can immediately distinguish between those objects that present themselves to us (occurring naturally) from the various artifacts that we create through our own activities and intentional behavior. This is captured in the qualia structure by reference to the Agentive role, whose default value of *nil* captures the primacy of a natural origin. Hence, the natural kinds *water*, *tiger*, and so forth, will have the following Agentive value:⁴

$$(92) \left[\begin{array}{l} \textit{water} \\ \text{QUALIA} = \left[\begin{array}{l} \text{F/C} = \mathbf{liquid} \\ \text{A} = \mathbf{nil} \end{array} \right] \end{array} \right]$$

By convention, when the value of a qualia role is *nil*, it will not appear as part of the qualia structure.

⁴It may be the case that specific aspects of how a natural object is brought into being can be represented in the Agentive, but we will not pursue this here. See also footnote 5.

Now consider how we can exploit this role to distinguish the different artifactual objects we encountered in the previous section, according to how they are created. Recall that nouns such as *cake* and *bread* were representing as denoting specific types of food, where the concept of *food* itself denotes a physical object with a Telic value making reference to the activity of eating. We can now make explicit reference to how an artifact is made by assigning that value to the Agentive, as illustrated in (94) for the noun *bread*.

$$(93) \left[\begin{array}{l} \textit{bread} \\ \text{QUALIA} = \left[\begin{array}{l} \text{F} = \mathbf{food} \\ \text{T} = \mathbf{eat} \\ \text{A} = \mathbf{bake} \end{array} \right] \end{array} \right]$$

For many artifacts, the Agentive value is recognized or identified well enough to be paraphrased with a single activity. For example,

- (94) a. [*house* [A= **build**]]
 b. [*painting* [A= **paint**]]
 c. [*letter* [A= **write**]]
 d. [*beer* [A= **brew**]]

For some nouns, however, we may be able to identify (classify) them as artifacts without knowing their exact provenance. This is most likely the case for most complex artifactual objects we encounter in our daily lives, such those denoted by the nouns *computer*, *car*, *refrigerator*, and so forth. In such cases, we identify to the best of our ability what the Agentive value should be. In these cases, a conventional speaker's lexicon would assign the underspecified activity of *make*, as shown below for *car*.

$$(95) \left[\begin{array}{l} \textit{car} \\ \text{QUALIA} = \left[\begin{array}{l} \text{F} = \mathbf{vehicle} \\ \text{C} = \{\mathbf{engine, door, wheels, ...}\} \\ \text{A} = \mathbf{make} \end{array} \right] \end{array} \right]$$

The empirical evidence suggesting that certain nouns encode information about the origin of the object is somewhat similar to the one we examined for the Telic in 2.2.3. For example, the contexts in (114) clearly evoke a creation act made available by *thesis*, *sentence*, *movie* and *painting*, even if this act is not expressed syntactically. That is, the sentences in (114) are interpreted as referring to specific events associated with the direct object, namely the events expressing how this object came into existence.

- (96) a. Paul completed his PhD thesis in 2000. (AGENTIVE = write)
 b. She wouldn't let me finish my sentence. (AGENTIVE = speak)
 c. Woody Allen has started a new movie. (AGENTIVE = direct, film)
 d. John began a large oil painting yesterday. (AGENTIVE = paint)

It is interesting to note that the verb *complete* tends to select the Agentive value of its complements (114), while the verb *finish* may select either the Agentive or the Telic value, depending on the context. For example, in the VP coordination in (98a), the presence of the verb *publish* seems to force an interpretation of the Agentive quale (*write*), while in (98b), the possessive *his* in the object NP appears to block an Agentive reading, hence allowing the selection of a Telic interpretation.

- (97) a. He just finished and published his first novel. (AGENTIVE = write)
 b. I have just finished his first novel. (TELIC = read)

As mentioned above, nouns with an Agentive value denote artifacts, objects made by humans for a particular purposes. This may be a concrete physical object, as in the examples above, or an abstract object, as in the case of the noun *idea*, which refers to a proposition (*prop*) that was brought about by an act of thought.

$$(98) \left[\begin{array}{l} \textit{idea} \\ \text{QUALIA} = \left[\begin{array}{l} \text{F} = \textbf{prop} \\ \text{A} = \textbf{think} \end{array} \right] \end{array} \right]$$

However, in GL, also action nominals like *arrival* and *building* and agentive nominals like *violinist* or *singer* have an Agentive value. We will return to agentive nominals below, while we discuss action nominals in the section on Nominalization in Chapter 12.

Objects with an identical Formal Qualia value may have different Agentive values, if they differ in the way they came into being and if this difference is encoded in the lexicon for those concepts. For example, consider liquids such as *water* and *coffee*, which share a Telic value reflecting that they are potable liquids, as illustrated in (100) below.

$$(99) \left[\begin{array}{l} \textit{water/coffee} \\ \text{QUALIA} = \left[\begin{array}{l} \text{F} = \textbf{liquid} \\ \text{T} = \textbf{drink} \end{array} \right] \end{array} \right]$$

It is clear that the general function of these objects, i.e., their Telic role, does not distinguish them conceptually. This is accomplished by an additional

dimension of classification, that of the Agentive quale, which specifies the mode of origin or creation (cf. (101)).

$$(100) \left[\begin{array}{l} \textit{coffee} \\ \text{QUALIA} = \left[\begin{array}{l} \text{F} = \mathbf{liquid} \\ \text{T} = \mathbf{drink} \\ \text{A} = \mathbf{brew} \end{array} \right] \end{array} \right]$$

While coffee is a liquid that is created typically by the activity of “brewing”, water is a natural kind which has been associated with a particular function. The identification of the Agentive role of “brewing” with the noun *coffee* acts to differentiate it from naturally occurring liquids. Notice how this is reflected in the interpretation of the Adjective-Noun construction in (102a) with *coffee*, as contrasted with that in (102b) with *water*.⁵

- (101) a. fresh coffee (AGENTIVE = brew)
b. fresh water (in contrast to “salt water”)

Similarly, the interpretation of the predicate *make* in the context of *cake* is ‘bake’, while in the context of *dress* it is ‘sew’. This suggests that the Agentive values of *cake* and *dress* are *bake* and *sew*, respectively.

- (102) a. Mary made a cake. (AGENTIVE = bake)
b. Her mother made her a dress. (AGENTIVE = sew)

The interpretation of the Agentive Quale differs from the interpretation of the Telic in one important respect. While the event in the Telic has a *generic* interpretation (as we saw in 2.2.3) in the sense that it refers to a typical situation, such an interpretation is not available for the Agentive quale. In fact, the occurrence of the event specified in the Agentive is a precondition for the existence of the entity. Technically, we say that the Agentive event is an *existentially bound* or *existentially quantified* event, that precedes the existence of the object. On this view, one of the defining properties for the Agentive quale is that it presupposes the occurrence of the event it encodes.

An exception to this would be the analysis proposed in Busa (1996) for “role-defining” agentive nominals like *violinist*. According to her analysis,

⁵There are natural kinds which do permit modification by the adjective *fresh*, such as “fresh milk”. Such cases involve an Agentive value that makes reference to activities that *enable* the object rather than *create* it. In this case, the activity of *milking* an animal. We will discuss these in Chapter 4.

a noun like *violinist* may be represented as denoting a person (F) having the ability (A) to play the violin (T). Under this analysis, the Agentive value of *violinist* encodes the precondition for the event specified in the Telic (i.e., the modal ability to play). The qualia structure representation proposed for the noun *violinist* is given below.

$$(103) \left[\begin{array}{l} \textit{violinist} \\ \text{QUALIA} = \left[\begin{array}{l} \text{F} = \mathbf{human} \\ \text{T} = \phi = \mathbf{play}(\mathbf{human}, \mathbf{violin}) \\ \text{A} = \mathbf{ability}(\phi) \end{array} \right] \end{array} \right]$$

Besides playing a role as an implicit predicate, there are a number of constructions in which the Agentive event associated with a noun and the noun itself co-occur in context. Again, there is an important distinction to be made between constructions in which the event being expressed corresponds to the lexically specified Agentive value of the noun and those where the event being expressed updates the Agentive information of the noun in composition.

Verb-argument constructions in which the verb expresses the Agentive value of the noun typically consist of a creation predicate and its direct object, such as those illustrated below in (105).

- (104) a. John baked a cake.
 b. They built a house in Greece.
 c. Mary took a photograph of her son.
 d. Marc is painting a picture.
 e. She wrote a letter to John.

An example of verb-argument construction in which the verb updates the Agentive value of the noun in composition is found in (106) where the predicate *type* expresses a manner of *writing*, that is, one of the possible ways in which the object *the letter*, is brought about, i.e., its Agentive value.

- (105) She sat down and typed the letter.

It is interesting to note that Adjective-Noun constructions are ungrammatical when the adjective expresses the Agentive value of the nominal head. For example, consider the semantically anomalous pairs in (107) below.

- (106) a. *baked bread (AGENTIVE = bake)
 b. *a build house (AGENTIVE = build)
 c. *a written book (AGENTIVE = write)

GL explains the anomaly of the examples in (107) as resulting from a lack of “informativeness” associated with the modification. Note that the expressions in (108) are anomalous because they are redundant:

- (107) a. *a male bachelor
b. *a female woman

Similarly, modification by adjectives denoting the Agentive value of the head are also uninformative. If, however, additional information is given besides reference to the Agentive activity alone, the resulting expression is well-formed, as illustrated in (109).

- (108) a. freshly baked bread
b. a well built house
c. a beautifully written book

Following this analysis, any semantic anomaly resulting from adjectival modification in Adjective-Noun constructions can be used as a diagnostic for identifying lexically specified Agentive values for nouns. Consider, for example, the expressions in (110):

- (109) a. Sue wore a knitted sweater to the party.
b. His later works include painted portraits.

While the modifying adjectives *knitted* and *painted* in (110) express how the objects named by the head nouns were created, the resulting expressions are semantically acceptable and do not appear uninformative. From this evidence, we may conclude that the adjectives in (110) do not identify the lexically specified Agentive values of *sweater* and *portrait*, respectively. That is, by identifying the mode of creation for the sweater as knitting, the expression “knitted sweater” is informative, since there are several ways in which a sweater may be made. Similarly, a portrait may be created by the act of painting, photography, or perhaps other means. The expression “painted portrait” is hence informative.

2.3 Conventionalized Attributes

As we have seen throughout this chapter, Qualia structure is intended to provide a systematic and linguistically grounded representation of aspects of word meaning that are usually not modeled in linguistic theory. There

are many properties and events, however, that are conventionally associated with an entity, but are not strictly part of the identified Qualia roles. In this section, we examine such information, which we will refer to as *conventionalized attributes* (CAs).⁶ A conventionalized attribute is a property typically associated with an object through our *experiencing*, by means of various perceptual modalities, rather than of our *use* of it (but see the observations in (120) below). Examples of CAs can be seen in (111):

- (110) a. Dogs bark.
 b. The sun warms the air.
 c. Water flows.
 d. Airplanes make noise.

The notion of recording the properties conventionally associated with an object is, of course, related to efforts of encoding our commonsense knowledge of things in the world (Lenat, 1989, Hobbs et al, 1985, Havasi et al, 200X). The motivation behind the notion of a CA is to provide a lexical means for encoding default information that can be used in semantic composition within the sentence, as well as for logical inferencing above the sentence.

CAs may be activated in specific contexts in a similar fashion to Qualia. For example, the contexts in (112) clearly evoke a *sound* event made available by the referent of the nouns *dog*, *bird*, and *rain*, while those in (113) require that the referents of the nouns *flower*, *gas*, and *coffee*, have the ability of creating a perceivable *smell*. Within GL, it is assumed that both these activities are encoded in the semantics of the nouns in the form of CAs.

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

Commonsense knowledge, such as that in (111a), provides the default values for contextualized interpretations, as in (112a).

It is a well-studied phenomenon that Qualia roles are accessed in certain contexts, such as those discussed earlier in the chapter, involving the

⁶See Pustejovsky, 2006, Pustejovsky and Jezek 2008, Pustejovsky and Rumshisky 2010 for first mention and development of this notion.

aspectual predicates *begin*, *start*, and *finish*. For example, the sentences in (114), repeated below, demonstrate how the Agentive role helps complete the interpretation.

- (113) a. Paul completed his PhD thesis in 2000. (AGENTIVE = write)
 b. She wouldn't let me finish my sentence. (AGENTIVE = speak)
 c. Woody Allen has started a new movie. (AGENTIVE = direct, film)
 d. John began a large oil painting yesterday. (AGENTIVE = paint)

But how can we account for the interpretations of the NP objects to the verb *enjoy* in (115)?

- (114) a. Mary sat out and enjoyed the sun. (warming up)
 b. It's a great place to enjoy the sea. (viewing, swimming, walking)

As with the aspectual predicates in (114), with the sentences in (115), the verb *enjoy* appears to reference an activity associated with the object, but not one that could be typically defined as filling the value of a Qualia role. Note that (115b) may invoke several activities made available by the noun's referent (*sea*-related activities). For example, as suggested above, it may evoke the activity of "viewing the sea", "swimming in the sea", "walking on the beach", and so forth. Within GL, it is assumed that the activities associated with *sea* may be classified according to a cline of conventionality, and that only the most conventional ones are coded in the noun's meaning as CAs. The details of the methodology adopted in GL to distinguish between coded vs. non-coded conventionalized activities will be presented in Ch. X. As we will see, this methodology is grounded on empirical evidence and focuses on distributional behavior as well as the comparison between elliptical and non-elliptical uses (for example, between "enjoy the sea" and "enjoy the sea view").

To see how CA interpretation differs from standard Qualia role values, consider the contextual interpretations we can associate with an adjective such as *fast*. When modifying artifactual nouns, it has long been observed that *fast* functions as an adverbial over the activity associated with the Telic role of the head noun it modifies.

- (115) a. Mary is a fast typist. (TELIC = type)
 b. This Porsche is a fast car. (TELIC = drive)

When there is no Telic role, however, the CA value(s) associated with the noun help provide an interpretation. Consider the sentences in (117) below.

- (116) a. The tuna is one of the fastest fish in the sea. (swimming)
 b. John was the fastest boy in the school. (running)

In GL it is assumed that conventionalized attributes are not external to Qualia Structure. Instead, they are interpreted as further characterizations of Qualia roles. That is, CAs are not seen as independent roles, but as projective manifestations of specific Qualia, which, together with the information coming from the Qualia, contribute to defining the overall semantic profile of a given type. The representation of CAs in GL is therefore always mediated through a Quale, as in (118) and (119) below.

$$(117) \left[\begin{array}{l} \textit{dog} \\ \text{QUALIA} = \left[\text{F} = \left[\begin{array}{l} \mathbf{animal} \\ \text{CA} = \mathbf{bark} \end{array} \right] \right] \end{array} \right]$$

$$(118) \left[\begin{array}{l} \textit{fish} \\ \text{QUALIA} = \left[\text{F} = \left[\begin{array}{l} \mathbf{animal} \\ \text{CA} = \{\mathbf{live_in(water), swim}\} \end{array} \right] \right] \end{array} \right]$$

That is, the commonsense properties of “barking” relating to dogs and “swimming” relating to fish, is encoded within the Formal role, but identified as specific CA.

Up to this point we have focused on CAs as associated with nouns denoting natural kinds, but CAs can associate with artifacts as well. For example the sentences in (120) clearly involve reference to a sound made by the objects denoted by the nouns *car* and *airplane*, and the most plausible interpretations for these sounds are indicated in parentheses.

- (119) a. I could hear a car behind me. (driving)
 b. We do occasionally hear an airplane. (flying)

The GL analysis of the context in (120a) may be viewed as follows: we experience the *sound* of cars through their use; hence, the noun’s Telic value *drive* interacts with the predicate *hear* to arrive at the conventionalized attribute for the car, i.e., the sound of it driving. The resulting representation is given below:

$$(120) \left[\begin{array}{l} \textit{car} \\ \text{QUALIA} = \left[\begin{array}{l} \text{F} = \mathbf{vehicle} \\ \text{T} = \left[\begin{array}{l} \mathbf{drive} \\ \text{CA} = \mathbf{make_noise} \end{array} \right] \end{array} \right] \end{array} \right]$$

2.4 Recursive Qualia Structures

Some scholars (Busa, 1996, Busa and Johnston, 1996, Bouillon, 1997, Pustejovsky, 1998) have suggested analyzing the individual Qualia roles as recursively making reference to qualia values themselves. One motivation for this comes from the interpretation of specific linguistic constructions including nominal compounds, where such recursive structure help capture the composition interpretations. For example, in the English compound *dining room* and the French compound *verre à vin* (wine glass), the Telic role for the entire expression is directly associated with “dining” and “wine”, respectively.

Now consider the following pair. The Telic role of the noun *rifle* is shown below as “firing”:

$$(121) \left[\begin{array}{l} \textit{rifle} \\ \text{QUALIA} = \left[\text{T} = \text{fire} \right] \end{array} \right]$$

Notice, however, that in a compound construction such as *hunting rifle*, the Telic value for the compound is no longer that of “firing” but is the more specialized activity of “hunting”.

$$(122) \left[\begin{array}{l} \textit{hunting rifle} \\ \text{QUALIA} = \left[\text{T} = \left[\begin{array}{l} \text{T} = \text{hunt} \\ \text{A} = \text{fire} \end{array} \right] \right] \end{array} \right]$$

In such cases, the Telic is seen as having a more complex structure; namely, by firing the rifle (the local Agentive value), it can be used for hunting (the local Telic value).

2.5 Qualia Structure for Verbs

Our discussion so far has focused on Qualia roles as they relate to noun meaning. Qualia structure may be extended, however, as a representational mechanism for describing all the major linguistic categories, including verbs and adjectives. It is easy to informally identify the Telic of a sandwich (*eating*), the Agentive of a cake (*baking*), or the Constitutive of bread dough (*flour*); it might seem less obvious, however, what the corresponding Qualia are, for verbs denoting such different situations such as *building*, *walking*, and *breathing*.

For the domain of individuals, the Formal Quale acts to identify the subdomain within which an entity is placed. That is, we know beer is a liquid, rocks are physical objects, jokes are abstract objects, and so on. Nominals refer to stable objects, and as such can be organized hierarchically with persistent concepts, denoted by the Formal Quale.

For the concepts denoted by verbs, however, things are more complicated. Consider again the questions addressed earlier in the chapter introducing the notion of the Qualia roles.

- (123) a. *Formal*: What kind of thing is it, what is its nature?
 b. *Constitutive*: What is it made of, what are its constituents?
 c. *Telic*: What is it for, how does it function?
 d. *Agentive*: How did it come into being, what brought it about?

Modifying these questions to the domain of verb denotations, we see immediately a correlation between distinct Qualia roles and the different eventualities denoted by verbs. For example, the Formal can be seen as characterizing predicates denoting stable and persistent verbal predicates, namely states such as *love* and *believe*. But since verbs can denote changes in the world, they can refer to the manner in which something happens or changes, that is, the Agentive Quale. For example, intentional activities such as those denoted by the verbs *run* and *walk* can be characterized as Agentive Quale 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). On the other hand, 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). Hence, we can briefly identify verbs through their Qualia structure as illustrated below with specific examples:

(124) STATE:

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

(125) ACTIVITY:

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

(126) CHANGE.STATE:

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

(127) ACCOMPLISHMENT:

$$\left[\begin{array}{l} \textit{build} \\ \text{QUALIA} = \left[\begin{array}{l} \text{T} = \mathbf{build.goal} \\ \text{A} = \mathbf{build.act} \end{array} \right] \end{array} \right]$$

In Chapter 4, we develop this model more fully, and show how Qualia structure interacts with event structure and argument realization.

2.6 Conclusion

Qualia Structure has predictive power with respect to the degree of conventionalization of certain activities. That is, activities that are related to the origin and purpose of an object are the most likely to be conventionalized and therefore included in a noun's lexical meaning.

Further Readings

Lexical representations based entirely or in part on qualia structure have been discussed and adopted in several other theoretical frameworks, including Jackendoff (1997, 2002), van Valin & Lapolla (1997), van Valin (2005, 2011), Cruse 2004 and Croft and Cruse 2004. From a computational perspective, Qualia have central in the work of Copestake (1993), Copestake and Briscoe (1991,1995), Bos, Mineur, Buitelaar (1993), Jensen and Vikner (1995), and the lexical database design and ontological architecture of the SIMPLE framework (Calzolari et al, 2000, Busa et al, 2001) and the Brandeis Semantic Ontology (Pustejovsky et al, 2006, Havasi et al, 2006).