Formal Ontology in Information Systems
R. Ferrario and W. Kuhn (Eds.)

doi:10.3233/978-1-61499-660-6-37

The Functions of Definitions in Ontologies

Selja SEPPÄLÄ [a,1], Alan RUTTENBERG [b], and Barry SMITH [c]
[a] *Department of Health Outcomes and Policy, University of Florida*
[b] *School of Dental Medicine, University at Buffalo*
[c] *Department of Philosophy, University at Buffalo*

Abstract. To understand what ontologies do through their definitions, we propose a theoretical explanation of the functions of definitions in ontologies backed by empirical neuropsychological studies. Our goal is to show how these functions should motivate (i) the systematic inclusion of definitions in ontologies and (ii) the adaptation of definition content and form to the specific context of use of ontologies.

Keywords. Ontology, definitions, textual definitions, logical definitions, functions of definitions

1. Introduction

What follows is a contribution to the understanding of definitions in ontologies based on a theoretical explanation of the functions realized by definitions when ontologies are developed and used. Our analysis draws on the more detailed account of terminological definitions provided in [1, 2]. We focus here on the lexically relevant components of ontologies, which we compare to lexical entries in terminological dictionaries.

We distinguish two main functions of definitions, cognitive and linguistic. We further identify specific functions of textual and logical definitions in ontologies. Our goal is to show how these functions should motivate (i) the systematic inclusion of definitions in ontologies and (ii) the adaptation of definition content and form to the specific context of use of ontologies.

2. Definitions in Ontologies

The term ‘definition’ refers to four types of things: a *cognitive activity* (DEF1) that produces a *cognitive representation* (DEF2), which constitutes the content of a *representational artifact* (DEF3) that concretizes this content and that is communicated in a *communication act* (DEF4). The four entities in our characterization may be defined as follows [3]:[2]

- **DEF1 – definition as a cognitive activity**: A cognitive activity performed by a cognitive subject that consists in forming a *cognitive representation* of some entity X and serves to specify what it is that makes an entity X rather than something else.

[1] Corresponding Author.
[2] We extend the work presented in Smith, *et al.* [4].

- **DEF2 – definition as a cognitive representation** [5, 4:59]: A cognitive representation composed of a set of items of knowledge or beliefs about a definition's object resulting from a *defining activity* in the sense of DEF1.
- **DEF3 – definition as a representational artifact** [4:59]: A representational artifact that expresses the *defining content* (DEF2) resulting from the *defining activity* (DEF1) and is communicated in a *defining act* (DEF4).
- **DEF4 – definition as an act of communication**: A communication act that consists of communicating *definition content* (DEF2) by means of a *representational artifact* to a receiver.

Our main focus is on the content and form of definitions (DEF2 and DEF3), which are the bearers of the analyzed functions.

2.1. Definition Contents

'Definition' in the sense DEF2 is a cognitive representation composed of features, i.e., pieces of knowledge/beliefs. Features represent properties of the object of the definition, things of a certain type or a particular thing in the world, that is the focus of domain-experts' activities and practices and to which they refer with specific term(s). These features form the contents of a definition in sense DEF2 that can be copied and concretized in various ways (DEF3).

Definitions are composed of at least two types of features: (i) the 'head' of the definition, called the *genus* (or *genus proximus* when it is the immediate superordinate type), which is used in an assertion of an *is_a* relation to the effect that the defined entity is a subtype of the type to which the genus refers; and (ii) one or more further parts, called the *differentiae*, that distinguish the defined type from other types of the same *genus*. Taken together, the genus and differentia(e) delimit the intension of the defined term (that which is said about the referent, i.e., a description of properties of the instances of the defined type), as well as its extension (the set of instances that fall under the intension). When definitions are viewed in these terms, we can distinguish four main logical forms:

Classical definition: A definition where the intension holds for all instances of the type X that is defined and does not hold for any instance that is not of that type. In this case, the characteristics expressed by the genus and differentia(e) are both individually necessary and jointly sufficient for something's being an X. This type of definition, which forms the ideal recommended case, is also called a definition by necessary and sufficient conditions. A standard example of classical definition is that of a *triangle* as: *A rectilinear figure that has three sides*. A classical definition of the species-genus form ('A is a B that Cs') is also called an 'Aristotelian definition', illustrated already in *A man is a rational animal*. Here *man* is the species, *animal* is the genus (or parent type) and *rational* is the specific difference — it is that feature of an instance of the genus which makes it also an instance of the species. Here too, the definition may have multiple differentiae.

Partial definition: A definition where the intension holds for all instances of the type that is defined but also holds for instances that are not of that type. A partial definition is a statement of necessary conditions that are not jointly sufficient. An example of partial definition is that of a *bird* as: *An animal that lays eggs*.

Typical or prototypical definition: A definition where the intension holds for most of the instances of the type that is defined, especially the typical ones, but also for instances that are not of that type. An example of a prototypical definition is that of a *swan* as: *An aquatic bird with a long neck, usually having white plumage*.

Instance definition or definite description [6]: A definition where the intension holds for only a single instance. This kind of definition would apply, for example, to resources that include what may be considered as proper names, such as the *Large Hadron Collider* (LHC) in an ontology of nuclear physics. In this case, the relevant kind of differentiae would probably inform us about the geographical location of the LHC and specify that it is (or was until some point in time) "the world's largest and most powerful particle accelerator."[3]

2.2. Definition Forms

The concretization form of the definition content depends on the context of use and the relevant (material) support on which the definition is to be communicated in that context (paper, electronic). The concretized representational artifact can be linguistic or nonlinguistic. In ontologies, definition contents are concretized in both forms for display on a computer screen and use by an automated reasoner: *textual definitions* are intended to be interpretable by human users, while *logical definitions*, in the form of axioms, are created, in part, so as to be able to be read and reasoned with by machines.

Moreover, each type of representational artifact has a surface form that depends on the target audience (including machines) and context of use. In its textual form, a definition consists, ideally, of a short sentence of the kind that is found in specialized terminological dictionaries. The surface form of the linguistic (textual) concretization of the definition content results from lexico-syntactic choices taken to meet human user needs. These needs can be linguistic, cognitive, and practical, such as, users' level of expertise, background, and the task being performed. Thus, the same defining content, for instance 'four legged', can be expressed as 'quadruped' when defining for specialists, and as 'that has four legs' when defining for laypersons.

Nonlinguistic concretizations include symbolic surface forms that also result from formalism choices adapted to the target audience and context of use (machine-readable logical formalisms and their different syntactic forms, first order logic, chemical symbols, etc.). The choice of formalism depends, for instance, on the formalism's syntactic and semantic properties, its expressivity, and the extent to which the formalism is known to the target audience. A widely used formal language for representing ontologies is the Web Ontology Language (OWL). In OWL, axioms function in a manner analogous to the necessary conditions previously discussed in relation to logical forms of definitions, where equivalence axioms correspond to classical definitions and subclass axioms to partial definitions.

Textual definitions and the axioms forming the logical definitions in ontologies go hand in hand in the ontology with other lexically relevant information, namely:[4]

- term (label, preferred label/name/term, synonym, etc.)
- unique identifier (IRI),
- indication of domain or scope,
- note or comment,
- example,

---

[3] Source: CERN, http://home.cern/topics/large-hadron-collider.

[4] As they can be visualized, for example, in the BioPortal tool, http://bioportal.bioontology.org. We only include lexically relevant information types, most of which are specified in OWL 'annotation properties'.

- illustration (graphic images, photographs).

These different types of information entity complement each other to provide the users with a specification of each defined term and its referent. This *complementarity principle* will be important to understand the cognitive function of definitions below.

3. Cognitive Function of Definitions

Definitions have primarily a cognitive function, that is, they produce some effect in the cognitive systems of their receivers. This function consists in reconfiguring and sometimes augmenting the receiver's beliefs in such a way as to fill the gap in knowledge that is implied by their definition consulting act [7:102]. Indeed, definitions provide knowledge and beliefs about the objects, processes, and so forth that are essential to the everyday activities and practices of domain-experts as reflected in their specialized vocabularies.

Definitions are consulted because it is presumed that the author of the definition has some knowledge about the term's meaning and properties of the term's referent that the receiver lacks [7:101-102, 112].[5] Any such definition consulting act should therefore involve some modification in the body of knowledge and beliefs of the receiver, including the sort of modification that consists in adding greater confidence to the receiver's beliefs.

To see why this is so, let us go back to the complementarity principle introduced in Section 2.2. If the receiver consults an ontology or a dictionary to look up a definition, then she presumably already possesses complementary information type(s), that is, a term or some perceptual knowledge of its referent, as shown in Figure 1. Therefore, the receiver's need lies at the semantic level of knowledge and beliefs.

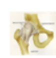
	What you are presumably looking for		
What you know	ligament	*Dense regular connective tissue connecting two or more adjacent skeletal elements or supporting an organ.*	
ligament		x	x
Dense regular connective tissue connecting two or more adjacent skeletal elements or supporting an organ.	x		x
	x	x	

Figure 1. Complementarity principle of information types corresponding to three states of knowledge: of term, of definition, of referent (or of an image of the referent).

The need on the part of the receiver that is to be satisfied by the definition can be either real or only pretended, as for example when a teacher pretends not to know the meanings of words when asking questions of her students. In standard uses of ontologies and dictionaries this need is presupposed. The definer is not directly acquainted with her

[5] That is, some knowledge about the intended meaning of a term by a group of competent speakers to refer to something that is the object of the definition.

receivers; but she can nonetheless be assured that the typical users of the artifacts she produces are in need of cognitive reconfiguration or augmentation of just this sort.

3.1. Theory of Lexical Competence

To put the above in terms of lexical competence, consulting an ontology or dictionary resource implies either (i) that the receiver's lexical competence is lacking (perhaps because it is marked by uncertainty and thus in need of confirmation), or (ii) that the receiver's lexical competence is to be modified to fit a specific context of use.

Marconi [8:2] divides *lexical competence* into two independent and cognitively motivated competences, one *inferential* (1) and one *referential* (2), further subdivided into *naming* and *application* abilities (2a and 2b). These correspond, respectively, to a speaker's ability to:

1. "have access to a network of connections between [a] word and other words and linguistic expressions", and thus perform different types of inferences;
2. a) name objects and circumstances in the world; that is, select "the right word in response to a given object or circumstance" (*naming*);
   b) apply a word to objects and circumstances in the world; that is, select "the right object or circumstance in response to a given word" (*application*).

These competences rely on two distinct systems:[6] the former on a semantic system; the latter on the perceptual and motor system. Neuropsychological studies show that one of these abilities can be lost or seriously damaged, while the other remains somewhat intact [9:132]. Both systems nevertheless interact and inferential competence plays a role in many referential performances, just as referential competences can enrich inferential competences.

Therefore, we propose, ontology elements and dictionary entries *adjust* receivers' lexical competence to converge towards that of competent speakers of a given domain (e.g., microbiology) in a given context of use (e.g., data annotation with an ontology of microbiology). Ontologies and dictionaries always have this cognitive function of lexical competence adjustment. Most if not all of the contents of the different fields of a dictionary entry contribute to the realization of this function, and something similar is true also in the case of ontologies (see Section 3.2).

To understand the specific cognitive function of definitions, we go back to Marconi's theory of lexical competence in [8:70-73], which puts forward the further hypothesis that "inferential competence may include several conceptually distinct and mentally separate abilities." He mentions, for instance, two specific sub-competences involving access to the *output lexicon* and the *semantic lexicon*.

- **Output lexicon:** "the words themselves (in either their phonological or graphic format)".
- **Semantic lexicon:** concepts or semantic representations "accessible from both words and pictures" and providing access to corresponding entities in the world.

We call the former *output inferential competence* and the latter, *semantic inferential competence*.[7] This subdivision is supported by cases of individuals capable of describing the properties and functions of an object without being able to name it [10:429, 11:132, 12:292 ff., 13:341, 342]. While Marconi does not pursue further this possible distinction, he nevertheless notes that "the ability to define words (word to definition) may be

[6] See Marconi [8:61-64, 141-142].
[7] Not to be confused with the semantic competence that is part of our general linguistic competence.

dissociated from the ability to find the word corresponding to a given verbal definition or to a description of the word's referent (definition to word)." [8:73]

The distinction is relevant for understanding what kind of cognitive mechanism or process lies behind the act of defining: defining (DEF4) is an act that goes directly or indirectly from word to definition. We schematize these distinctions in Figure 2.

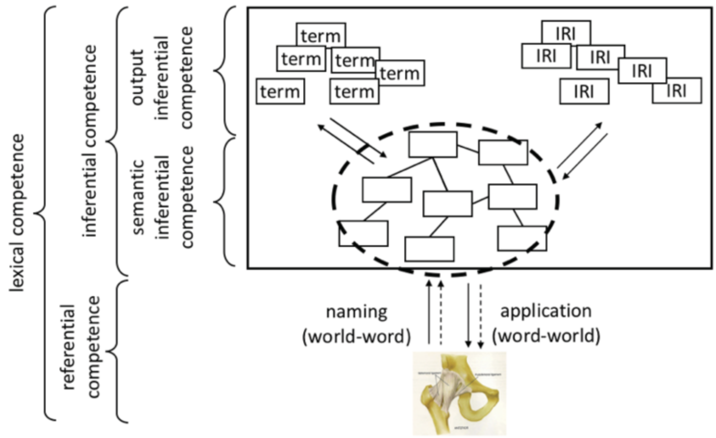


Figure 2. Lexical competence and its sub-competences (based on [14] and [8]).

Following these subdivisions, we propose that a definition adjusts the overall lexical competence of its receivers by adjusting their inferential competence, and, more specifically, their semantic inferential competence. The effect on referential competence is realized indirectly, for example, through those words in the definition in relation to which the receiver already enjoys referential competence.

3.2. Complementary Information Types for Complementary Adjustments

Based on Marconi's intuition "that the two sides of lexical competence, inferential and referential, mostly rely upon different kinds of information" [14:149] and considering the above subdivision of inferential competence, each information type in a dictionary entry and ontology element can be paired to a sub-type of competence. To see this, we associate the three types of lexical sub-competences introduced above[8] (referential competence, semantic inferential competence, output inferential competence) with the lexically relevant complementary information types in an ontology element introduced in Section 2.2. These pairings are illustrated in Figure 3.

Referential competence: Ontologies can include *examples* among their elements; and also, though more rarely, *illustrations* (for example photographic images). Both allow users to recognize a described portion of reality on the basis of exposure to information types related to referential competence. In the case of examples, the example text involved will standardly call in aid only the user's referential competence — thus it will not explicitly convey any information pertaining to the properties of the referent of the sort that would be provided by a definition. Illustrations can be provided by the

[8] This parallel between different types of competences and dictionaries is sketched several times by Marconi [8:56, 66-67, 114, 146, 157]. In [1], we extended it to dictionary entry fields. Here, we apply the proposed pairings to ontologies.

ontology developers in the form of links to images accessible through the ontology interface[9] or by the ontology's users through the annotation of images with ontology classes, for example, when images of plant formations are annotated with the Plant Ontology (PO) [15].

Semantic inferential competence: Those information types in an ontology that relate to semantic inferential competences are the *textual definition* and *axioms*, the *notes* (in 'comment' annotation properties), and an *indication of domain* or *scope* of the ontology. These parts of an ontology element provide information that contributes to a user's understanding of the intended meaning of the ontology classes and their use in inferential, i.e., logical, operations. They convey the body of knowledge and beliefs that can reasonably be considered to be shared by both ontologists and domain experts when they use a given term.

Output inferential competence: The information types that relate to this competence are the labels (a word or a phrase), including not only preferred label and synonyms but also the IRIs identifying classes or relations. IRIs are included under this heading insofar as ontologies are manipulated by machines and the IRIs are the symbolic forms that a machine uses.

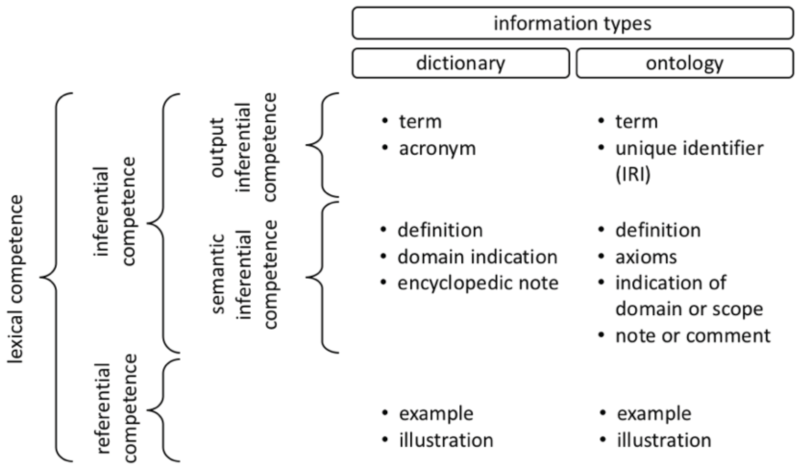


Figure 3: Pairings of information types in a dictionary entry and ontology element with the sub-competences in which they are involved.

Thus, if we consider that each type of lexically relevant information in an ontology element and dictionary entry is involved in one of the three lexical competences (referential, output inferential, semantic inferential), then we can say that definitions have primarily the function of adjusting receiver's semantic inferential competence — and, thus, of influencing their beliefs and reasoning.

The cognitive function of definitions, in sum, is to bring about a belief reconfiguration or belief augmentation on the part of the receiver such that it allows her to competently use the defined term in inferential processes and, indirectly, also in referential ones. For this adjustment to be successful, the definition has to be relevant with respect to a specific target audience and context of use. We address definition

[9] See, for example, the Foundational Model of Anatomy (FMA) ontology browser at http://xiphoid.biostr.washington.edu/fma/index.html.

relevance in the context of ontologies in the next section, in relation to the linguistic function of definitions.

4. Linguistic Function of Definitions

In most cases, consulting an ontology or dictionary reflects a need to align oneself with a certain pre-existing lexical use. Even though consulting such resources does not necessarily answer a lexical question as such, the fact that both ontologies and dictionaries include lexical units implies, at least indirectly, that they fulfill a linguistic function.

Someone consulting an ontology or dictionary, and thus a definition, aims to attune their lexical competence in order to promote their use of given linguistic signs in a way that converges towards that of competent speakers in specific sorts of contexts.[10] In ontologies and terminological dictionaries, these competent speakers are either (i) domain experts, where the resource in question covers domain-specific terms, or (ii) ontology or terminology experts, for matters pertaining to domain-independent ontology or terminology terms.

A single lexical unit (graphic or phonic sign) can have different semantic values depending on the context of use, for example, on the domain of expertise (e.g., banking, nephrology) and the task at hand (e.g., understanding a text, annotating scientific data). Semantic value is thus determined in part by the context of use of given competent speakers. Definitions adjust receivers' lexical competence toward the corresponding global norm of use, that is, a semantic value acknowledged by a given speaker community.

Yet, definitions are also often used stipulatively [17, 18], to describe or prescribe a use that is not the global norm. The effect of the definer creating the definition is to establish local convergences (between themself and successive readers) to avoid confusion with and ambiguities within the global convergent norm of use.

Both global and local uses can appear in free discourse and text that constitute the basis for the defining activity encapsulated in ontologies and dictionaries. Ontologists and dictionary authors are, after all, mediators — they usually do not create norms of use but rather mediate between, for instance, authors of scientific textbooks and the users of terminologies who require support from definitions. Their secondary defining activity consists in making explicit the meanings underlying uses of terms of distinct provenance: either a global norm of use or a local use.

Once a definition is included in an ontology or dictionary, it thenceforth expresses a lexical use, which is more or less stipulative (prescriptive) depending on the context of use of the ontology or dictionary, the intention of the definer, the attitude of the receiver, and so on. A definition can thus be regarded as being on a scale that ranges from describing a use (descriptive definition) to having a regulatory or normative function on use (stipulative or prescriptive definition).

Although an empirical question, it is reasonable to assume that the more a definition is descriptive and aimed at conveying a term's meaning, for example, for understanding a text, the more it is likely that it will have a (proto)typical logical form, with a combination of necessary and (proto)typical features. On the contrary, the more a

[10] Throughout this section, we use Marconi's theory of Lexical Competence [8]. For a similar pragmatic approach to competence adjustment and to its directionality (influenced by John Searle), see [16:100-101].

definition aims at stipulating a meaning, the more it tends to include only necessary (and possibly sufficient) conditions, which allow to disambiguate the defined term.

The relationships between types of resources, their lexical uses, and the logical forms, utility, and cognitive effect of their definitions is illustrated in Table 1. The table shows the spectrum of properties of more lexically-oriented resources on the left-hand side and ontologically-oriented resources on the right-hand side. The former tend to have more descriptive and (proto)typical definitions, employed mostly to convey meaning. The latter tend to include more prescriptive definitions with necessary and, possibly, sufficient conditions, employed mostly to disambiguate meaning. The corresponding cognitive effects are, respectively, belief-augmentation and belief-reconfiguration. Note that whenever a definition fulfills a linguistic function a cognitive function is also fulfilled — but that the reverse does not hold.

Table 1. Relationships between types of resources, their lexical uses, and the logical form, utility, and cognitive effect of their definitions.

	Type of resource	
	Lexicographic, lexico-semantic, terminological	**Ontologies, standards, normative terminologies**
Lexical use	describe	prescribe
Logical form of definitions	(proto)typical	necessary & sufficient conditions
Utility	conveying meaning	disambiguating meaning
Cognitive effect	knowledge/belief augmenting	knowledge/belief reconfiguring

In sum, the linguistic function of a definition is to convey or disambiguate the semantic value of a term in a more or less descriptive or stipulative way, by delimiting its intension and extension by means of more or less (proto)typical and classical definitions. The resulting cognitive effects on the body of knowledge and beliefs of the receiver are more or less to augment and reconfigure them.

5. Functions of Definitions in Ontologies

5.1. Stipulation of the Intended Meaning of a Term

In ontologies, the represented knowledge is not always related to a lexical unit that is naturally used by speakers. The represented classes are usually labeled with terms from a controlled vocabulary with an intended meaning, where terms from the vocabulary may or may not be used in natural-language contexts. Whether an ontology term is commonly used by domain experts or not, its definition specifies the intended meaning of the term in a normative way. They allow ontology users to use these terms in a competent manner; they enhance users' lexical competence with respect to ontology terms by providing definitions that clarify and disambiguate.

The corresponding adjustment of semantic inferential competences is, therefore, mostly a belief-reconfiguring one — as opposed to a belief-augmenting adjustment. When consulting a definition in an ontology, the receiver (usually a domain-expert) has to set aside the body of knowledge accumulated over the years about the terms of her domain and their referents, and restrict the meaning of the term to only those items of knowledge intended by the ontology developers. This involves some reconfiguration in her beliefs.

5.2. Term Disambiguation

Ontologies and controlled vocabularies aim at aligning the lexical use of their users to achieve intra- and inter-personal consistency, for example, when annotating scientific data or integrating databases with an ontology. Correspondingly, definitions fulfill a stipulative linguistic function; that of adjusting the receiver's lexical competence towards a local use. In order for a definition to realize this function, it has to be tailored in such a way that its logical form leaves no room for ambiguity. It therefore has a disambiguating function.

To realize this function, ontologies would ideally provide Aristotelian definitions in which the genus (more precisely, the immediate superordinate category or *genus proximus*) is specified along with the differentiae to provide a statement of individually necessary and jointly sufficient conditions that precisely distinguishes the intended meaning of the term from that of neighboring terms. Provision of such Aristotelian definitions is costly, however, and in some cases it is not possible at all because of lack of scientific knowledge; thus ontologies (especially large ontologies) often make do merely with the statement of necessary condition as tool for disambiguation.

Where definitions are provided using a formal language such as OWL, they take the form of axioms which serve to disambiguate terms in a way that is analogous to the way textual definitions serve this purpose. Every subclass axiom represents a necessary condition that all instances in the extension of the term need to satisfy. These axioms serve to determine the extension of a term by restricting it to those entities meeting the asserted condition. Each additional axiom restricts the extension further, though it is in many cases not possible to restrict the term to only its intended extension by providing conditions that are jointly sufficient

For the most part, a class (as opposed to a class expression) serving as relatum in a subclass and equivalence axiom should correspond directly to the genus in the textual definition, as in the definition of *bacteremia* (Figure 4). The other defining conditions are expressed by non-atomic class expressions.

```
unblinding process
SubClassOf
        planned process
        (part_of some study design execution)
          and (part_of some informing subject of study arm)
```

Here, two axioms define the term *unblinding process* in the Ontology for Biomedical Investigations (OBI). 'Planned process' is the asserted superclass of 'unblinding process', as in: 'unblinding process is_a planned process'. 'Part_of some study design execution' and 'part_of some informing subject of study arm' are class expressions that in conjunction restrict the extension of 'unbinding process' to only those planned processes that are part of a study design execution and that inform the subjects about the study arm in which they participate.

6. Functions of Logical Definitions

In addition to the cognitive and linguistic functions described above, which apply both to textual and logical definitions, we distinguish three primary functions of logical definitions: *instance classification and consistency checking*, *taxonomic schematization,* and *regularizing expression of facts*.

6.1. Instance Classification and Consistency Checking

Classical definitions function in instance classification and consistency checking. Necessary conditions serve as checklists for determining whether an instance is consistent with the classes of which it is asserted to be a member. When an instance's properties are consistent with sufficient conditions for a class, that instance can be asserted to be of that class. Indeed, the linguistic function of a definition is to convey the semantic value of a term by delimiting its intension and extension. Since ontologies provide definitions with only necessary conditions to delimit the intension of a term, their definitions can be used as reliable heuristics for such classification tasks.

This function is further useful for identifying errors in definitions, since it allows the definer to test a definition's scope by seeing whether it classifies the right instances and is able to exclude, as inconsistent, unwanted ones. The definition content may thus be checked and corrected to ensure that it is about the definition object and only the definition object. For example, if there are exemplars of the definition object that don't concord with the definition then the defining content is adjusted so as to include those. Similarly, if there are entities that concord with the definition, but are not exemplars of the intended definition object, then the defining content is adjusted to exclude them.

6.2. Taxonomic Schematization

We call the first function 'taxonomic schematization'. When employed in this capacity, the logical definition of a class provides a schema or template for the axioms of its subclasses. The goal is to provide robust, principled taxonomic relations between parent, child, and sibling classes. The axioms specified for each class are true of all its subclasses. This makes it possible to use axioms to specify differentiae for its child classes, in other words, to use these axioms as templates for the axiom of the subclasses, as well as for the contents of the associated textual definitions [1, 2].

This can be done by asserting a relational axiom for the parent class relating it to some other kind of entity (e.g. by writing an axiom for a class X asserting that any X part_of some Y). For every subclass of this related kind, a subclass of the parent can then be distinguished. The property relating both relata can also be a sub-property of the more general one. For example, the axiom specifying the term *infection* in the Infectious Disease Ontology (IDO):

```
infection
SubClassOf material entity
    (has_part some infectious agent)
       and (part_of some extended organism)
```

can be used to generate the subclass axioms of its child terms, such as *bacteremia* (Figure 4).

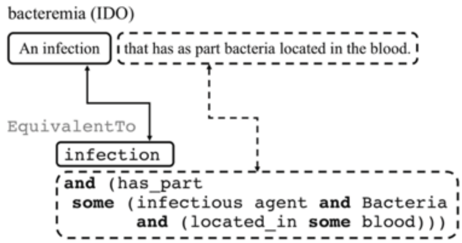


Figure 4. Correspondences in the parts of the textual definition and the axioms of the IDO term *bacteremia*.

6.3. Regularizing Expression of Facts

An ontology can be considered a specification of a controlled vocabulary for expressing facts in a given domain. The controlled vocabulary used in axioms allows the user to check the intended meaning of a natural language expression in the textual definition. However, as noted by Stevens, *et al.* [19:6], textual definitions are not simple sentence-by-sentence verbalizations of axioms. The axioms create unordered lists of sentences that can involve redundancies. A textual definition is rather a grouping of one or more axioms that form a non-redundant, fluent paragraph. The expressions used in natural language definitions are thus more idiomatic. Expressions such as 'continuant_part_of' or 'inheres_in' are after all, not very natural.

Such a vocabulary is also much sparser than the vocabulary that would be used to express these facts in natural language, that is, there is a one-many correspondence between ontology terms and words in domain-relevant portions of natural language. This means that the syntax for expressing facts (i.e., assertions between instances) using ontology terms necessarily diverges from the syntax used for expressing the same facts in natural language.

An important function of axioms in ontologies is to provide a schematic indication of how this should be done. Axioms complement textual definitions in contributing cognitively towards regularizing users' use of terms. Consider, for example, the class expression '*develops_from* some hematopoietic stem cell' as it occurs in one of the axioms involving the term 'leukocyte' in the Cell Type Ontology (CL) (Figure 5).

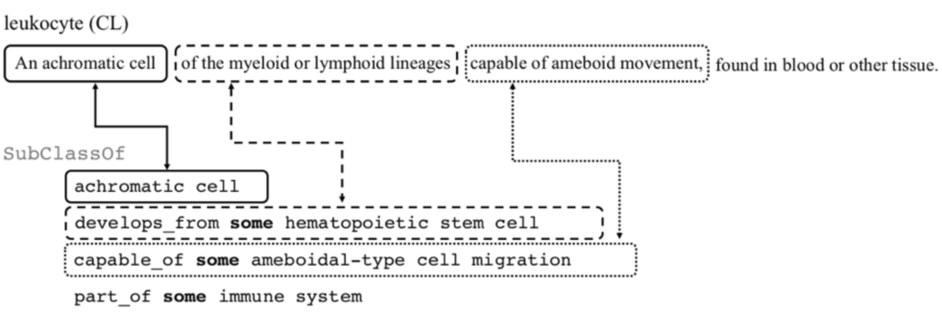


Figure 5. Correspondences in the parts of the textual definition and the axioms of the CL term *leukocyte*.

In this expression, the relation '*of the … lineages*' in natural language is expressed at the logical level by the 'develops_from' relation that is part of the controlled vocabulary of the ontology.

7. Conclusion

Our goal was to further our understanding of what ontologies do through their definitions to motivate (i) the systematic inclusion of definitions in ontology resources, and (ii) the adaptation of the content and form of definitions to serve the specific needs of ontologies. We extended theoretical suggestions of Marconi on lexical competence by introducing a subdivision of inferential competence into *output inferential competence* and *semantic inferential competence*. We also built on another suggestion by Marconi to propose a pairing of different sub-competences to linguistically relevant information types in ontology elements.

Against this background, we showed that the cognitive function of definitions is to adjust (reconfigure or augment) receivers' overall lexical competence by adjusting their inferential competence, and, more specifically, their semantic inferential competence in such a way as to become more closely aligned to the knowledge or beliefs of relevant experts. Our pairings show that terms and IRIs alone do not fulfill this function – although they might indirectly for human users, provided they enjoy the appropriate lexical competence. Of course, a minimal taxonomic hierarchy allows the receiver (human or machine) to engage to some degree in semantic inferential processes, but in a very limited way. Our pairings also show that other information types, such as comments or indications of domain or scope, also participate in the adjustment of a receivers' semantic inferential competence. While these information types may be useful for human users (at the risk of introducing ambiguity), they are not used by automated reasoners. Therefore, ontologies should systematically include both textual and logical definitions.

We showed, further, that this cognitive adjustment is closely related to a definition's linguistic function, whereby belief-reconfiguration or belief-augmentation results in a convergence in the use of a linguistic sign on the salient use of that sign by the community of experts. For that convergence to take place in the specified context of ontologies, the content and form of the definitions must be adapted correspondingly. The particular context of use of ontologies gives precedence to the logical definition, since it is the one used by the reasoner to perform logical operations. If the reasoner fails due to a poor definition, the main purpose of the ontology is defeated. Therefore, the contents of the definition have to be adapted for use by machines. One such adaptation is to include necessary, and whenever possible, sufficient conditions.

Yet, stating only necessary (and sufficient) conditions in a textual definition might be too limited for an adequate understanding of the defined term. It may be useful to add extra information about the defined term's referent, such as typical features. But such non-necessary conditions could, in some logics, such as description logics, make reasoning more difficult. Thus, any extra information that might be useful for human understanding should be included, for such logics, in the form of a comment that may be complemented with examples.[11] As we saw, all these information types complement each other to enhance an ontology user's overall lexical competence; the definition adjusts it to a specific context of use, which in ontologies requires disambiguating terms. The complementary information must however be controlled to avoid introducing ambiguity. Furthermore, in order to ensure consistency in ontology development and use,

[11] Another example is provided by the NEON ontology annotation properties, which include "two types of definitions "stored" within one definition value" [20:3]: a *textual definition* and an *ontological definition*. While both use natural language, the former is expressed in free-form text and the latter is limited to restating the information in the axioms.

the textual and logical definitions of a term must convey the same type of content. Thus, axioms can be used as content templates for lower-level categories and for textual definitions (provided the axioms are complete).

Acknowledgement

Work on this paper was supported in part by the Swiss National Science Foundation (SNSF). We also thank Peter Elkin of the University at Buffalo for invaluable assistance and the anonymous reviewers for their useful comments.

References

[1] S. Seppälä, *Contraintes sur la sélection des informations dans les définitions terminographiques: vers des modèles relationnels génériques pertinents*. PhD Thesis, Université de Genève, 2012.
[2] S. Seppälä, An ontological framework for modeling the contents of definitions. *Terminology*, **21**(1): 23–50, 2015.
[3] S. Seppälä, A. Ruttenberg, Y. Schreiber, and B. Smith, Definitions in Ontologies. *Cahiers de Lexicologie*, forthcoming.
[4] B. Smith, W. Kusnierczyk, D. Schober, and W. Ceusters. Towards a reference terminology for ontology research and development in the biomedical domain. In *Proceedings of KR-MED*, 2006.
[5] B. Smith and W. Ceusters. Aboutness: Towards Foundations for the Information Artifact Ontology. In *Proceedings of the Sixth International Conference on Biomedical Ontology (ICBO 2015)*, 2015.
[6] B. Russell, On Denoting. *Mind*, **14**(56): 479–493, 1905.
[7] J. Sager. *A practical course in terminology processing*. John Benjamins, 1990.
[8] D. Marconi. *Lexical Competence*. Language, Speech and Communication Series, a Bradford Book. The MIT Press, 1997.
[9] M.J. Riddoch and G.W. Humphreys. Visual object processing in optic aphasia: A case of semantic access agnosia. *Cognitive Neuropsychology*, **4**(2): 131–185, 1987.
[10] R.A. McCarthy and E.K. Warrington. Evidence for modality-specific meaning systems in the brain. *Nature*, **334**(6181): 428–430, 1988.
[11] M.J. Riddoch and G.W. Humphreys. A Case Of Integrative Visual Agnosia. *Brain*, **110**(6): 1431–1462, 1987.
[12] T. Shallice, *From neuropsychology to mental structure*. Cambridge University Press, 1988.
[13] E.K. Warrington. Agnosia: The Impairment of Object Recognition. *Handbook of Clinical Neurology*, **45**: 333–349, 1985.
[14] D. Marconi. On the Structure of Lexical Competence. *Proceedings of the Aristotelian Society*, **95**: 131–150, 1995.
[15] N.T. Lingutla, J. Preece, S. Todorovic, L. Cooper, L. Moore, and P. Jaiswal. AISO: Annotation of Image Segments with Ontologies. *Journal of Biomedical Semantics*, **5**: 50, 2014.
[16] M. Riegel. La définition, acte du langage ordinaire: De la forme aux interprétations. In Centre d'études du lexique, editor, *La définition*, Librairie Larousse, pages 97–110, 1990.
[17] K. Ajdukiewicz. Definitions. In *Pragmatic Logic*, Springer Netherlands, pages 57–84, 1974.
[18] A. Gupta. Definitions. In E.N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*, 2015.
[19] R. Stevens, J. Malone, S. Williams, R. Power, and A. Third. Automating generation of textual class definitions from OWL to English. *Journal of Biomedical Semantics*, **2**(Suppl 2): S5, 2011.
[20] G. Frishkoff, P. LePendu, and S. Nikolic. *NEMO Ontology Annotation Properties: Ontology- and Concept-level Properties for Annotation*. NEMO Technical Document 2010-002, 2010.