JoSS Journal of Speech Sciences

DOI: 10.20396/joss.v14i00.20396

TOWARD A FORMAL ONTOLOGY FOR CORPUS LINGUISTICS

ALMEIDA, Mauricio[1*]
RASO, Tommaso[2]
MELLO, Heliana[3]

[1] Escola de Ciência da Informação, Universidade Federal do Estado de Minas Gerais - ORCID: https://orcid.org/0000-0002-4711-270X
[2] Faculdade de Letras, Universidade Federal do Estado de Minas Gerais – ORCID: https://orcid.org/0000-0002-3446-313X
[3] Faculdade de Letras, Universidade Federal do Estado de Minas Gerais – ORCID: https://orcid.org/0000-0003-0267-9005

Abstract: *Ontologies have been among us for about 30 years, proving their importance in domains like medicine, biology, and law, to mention a few. They are studied in multiple scientific disciplines, including philosophy, computer science, information science, and even linguistics. In this paper, we present an initiative to build an ontology for corpus linguistics in the scope of the C-ORAL-BRASIL project. We first introduce the interdisciplinary nature of ontologies, asking what they are, both in linguistics and fields of artificial intelligence, like knowledge representation. Then, we explain the top and middle levels that integrate the initiative and present a fragment of the corpus linguistics ontology, focusing on more relevant classes. Definitions, labels and relationships of such fragments are presented in the appendix. The complete resulting ontology counts around 260 classes and 90 relationships distributed over three layers: domain, middle and top-level ontologies. Finally, we offer our final remarks about the experience.*

Keywords: Ontology; Corpus; Focus; Language.

[*] Corresponding author: ba@eci.ufmg.br

1 Introduction

Questions about ontologies and their uses have existed for over thirty years. Perhaps the most famous attempt at an answer comes from Thomas Gruber's 1992 article "What is an Ontology?"[1] Gruber is a former Stanford University professor whose research in ontology is responsible for the SIRI technology, a virtual assistant built into virtually all Apple devices (Almeida, 2020). The first task in researching ontology is distinguishing "ontology" and "ontologies." This distinction makes the term's meaning precise in the specific context of the research.

Initial use of the term "Ontology" refers to the discipline of Philosophy that deals with the existence and most general forms of that which exists. The term refers to Aristotle's studies but only appeared in texts by German philosophers in the 17th century. We here call this approach "ontology as a discipline."

However, the term "ontologies" has taken on other meanings beyond philosophy. A distinct meaning has emerged in applied scientific fields, especially those that manipulate data, information and knowledge. In the scope of such applied sciences, ontologies refer to representational artifacts. These artifacts, inserted into information systems, provide a structure of the world to machines, allowing them to perform a basic type of reasoning. We call this approach "ontology as artifact."

Combining these two approaches—ontology as a discipline and artifact—we find the so-called "applied ontology" (Smith, 2004; Guarino, 1998). Thus, in applied ontology, philosophical principles are used to create well-founded models of reality to be employed in ontologies as artifacts. Ontologies as artifacts, in turn, compose information systems.

This paper is about applied ontology and presents an ongoing initiative for developing an ontology as artifact for corpus linguistics. We introduce the term "ontology" in some fields to better characterize the research in ontologies (Section 2). Then, we describe the Informal Spoken Brazilian Portuguese Reference Corpus (C-ORAL-BRASIL) Ontology, built through the expertise of linguists (Section 3). Finally, we present our final remarks (Section 4).

2 What is an ontology?

In this section, we introduce views of ontology and ontologies in some research fields: philosophy (Section 2.1), Computer Science, mainly in modeling and knowledge representation (Section 2.2) and linguistics (Section 2.3).

2.1 Ontology in philosophy: ontology as a discipline

Metaphysics is the branch of philosophy concerned with the study of reality. Ontology is a fundamental discipline that occupies a prominent position within metaphysics. In the 18th century, German philosophers used the term "ontology" to refer to a discipline that included notions of being, quality and quantity, and truth and falsehood. In the 20th century, the term was used in philosophical discussions as another name for a scientific logical system (Hennig, 2008). The present section deals with ontology as a discipline.

There is general agreement that ontology within Philosophy is concerned with what kinds of things can exist. Here, "kind" means category, a term that Aristotle used to discuss the kinds of statements that could be made about an object. A theory of categories is the most important topic of any ontology. Such theories specify category systems that are structured in hierarchical levels. Sometimes, these systems have the form of an inverted tree in which the topmost category is named entity. In this case, anything may be described as an entity of some sort, but the next

[1] Available on the Internet: https://tomgruber.org/writing/definition-of-ontology/

level of categorization is controversial. This rationale applies to the abovementioned case, in which only a unique topmost ontology exists, but whether there is such a unique category is still a matter of debate. Moreover, other difficulties arise in discussing systems of categories, for example: How many categories exist? How to discern category levels?

At least three systems influence contemporary ontological studies: the Aristotelian, Kantian, and Husserlian systems (Thomason, 2009).

The Aristotelian tradition provides a method for ordering categories of things based on their essential characteristics. Although Aristotle's method for ordering is well known, it leaves open the question of the ontological status of real essence. Kant also posed a question concerning the unsystematic way that Aristotle's categories might have been chosen, making it impossible to know whether Aristotle's list of categories is complete (Jansen, 2008).

Although it deals with the categories of human understanding, Kant's theory retains ontological importance as the categories apply *a priori* to all objects of cognition. By delineating the concepts necessary for cognition, one can acquire knowledge of the categories governing any object of cognition but not the things themselves.

Husserl conceived a formal ontology discipline analogous to the previously developed "formal logics" (Johansson, 2004). Formal disciplines are distinct from material ones in that the former apply to all domains of objects and are independent of the peculiarities of a specific field of knowledge.

Although these positions are not always clearly stated, one can see that, whereas Aristotle was guided by ontological realism, Kant and Husserl subscribed to some form of antirealism. Kant, in general, is considered an idealist, asserting that reality is fundamentally mental, mentally constructed, or otherwise immaterial. Husserl is said to adhere to versions of both realism and idealism (Beyer, 2011).

The term "realism" has several interpretations in Philosophy, but the main aspects that permeate it are the existence and independence of the mind. The question of existence refers to how entities exist by the facts of reality. For example, the temperature of the Sun is independent of linguistic practices or human interventions. In realism, universals explain the psychological capacity to recognize, group, and classify particulars – the so-called Problem of Universals. Both universals and particulars are notions that date back to Aristotle and Plato, meaning, roughly, groups of things and things in themselves (Almeida, 2020).

A universal corresponds to a set of repeatable and exemplifiable characteristics only when instantiated in their respective particular. In the context of science, universals represent abstract entities named by generic terms and used to formulate scientific laws. A particular is an entity that inhabits reality, is not repeatable, and remains a unique instance of a universal. Each particular maintains a relationship of instantiation with a universal and occupies a unique space-time location (MacLeod & Rubenstein, 2005).

Any solution to the Problem of Universals that is different in any respect from the particular-universal solution is identified as anti-realism. It is easy to be an anti-realist because refuting the entire thesis is unnecessary, only a single premise. Common types of anti-realism are "nominalism" and "conceptualism."

For the nominalist, the universal does not exist; if it did, determining and obtaining knowledge about it would not be possible. Nominalists argue that the Problem of Universals can be solved by reflecting on particulars. The nominalist critique also suggests that characteristic similarities are not decisive for categories. The first theories of nominalism assumed a linguistic character, establishing that knowledge would be in particulars because it is about them that people make abstractions. If "something" in common exists between entities, this something could exist only in the mind and be expressed through words or signs (Rodriguez-Pereyra. 2015).

Conceptualism emerges as an alternative to explain the Problem of Universals. It seeks to do so by sharing something individuals would have in common, as long as that does not involve universals. For this very reason, "concepts" are adopted. Concepts, unlike universals, are mental entities (Almeida, 2020).

2.2 Ontology in Computer Science: the ontology as an artifact

The present section deals with ontologies as artifacts. Ontologies as artifacts have their origin in the semantic networks of the 1960s. The term "semantic network" arises from Ross Quillian's[2] Ph.D. thesis in which it was used as a network of information models. The term came to be applied for different purposes, to model all sorts of non-semantic things such as propositions in logic, the physical structure of objects, and the behavior of devices. In artificial intelligence and knowledge representation, a semantic network generally refers to an extensive class of informal and formal symbolic representations (Stefik, 1995). These representations have in common that they are all made of links and nodes (Figure 1).

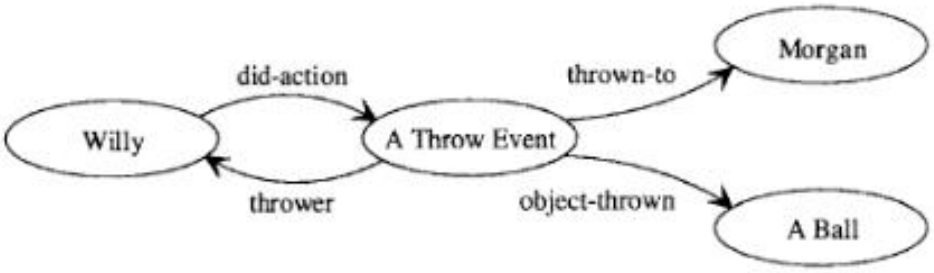

Figure 1: Example of a semantic network
Source: Stefik (1995)

In the 1990s, the term "ontology" continued to be used within the set of technologies encompassed under the label "semantic web," which brought a renaissance to artificial intelligence by promising to bring the possibility of automatic inferences to the web.

Even considering only computer science, the study of ontologies as artifacts is characterized by the coexistence of diverse definitions (Guizzardi, 2005). Smith (2004) distinguishes two types of ontologies to address the issue of term usage plurality. The first relates to how the universe is organized and corresponds to the philosophical approach, and the second concerns conceptualizing a domain. This second fulfills the need to express an ontology as an artifact within the scope of software engineering and knowledge representation.

So, the term ontology has two primary meanings in computer science. The first concerns using ontological principles to understand and model reality. This aligns with its original role in philosophy, namely, to provide an account of what exists which began to be used to characterize entities in modeling activities. The second concerns representing a domain in a knowledge representation language. Ontologies as artifacts consist of statements written in this language, which automatic reasoners can process.

An example of the first approach is the work on evaluating models according to ontological principles (Wand & Weber, 1990; Wand, Storey, & Weber, 1999). An example of this second approach is using the web-oriented knowledge representation language, OWL (Web Ontology Language), to build ontologies within the semantic web context (Staab & Studer, 2004).

[2] An important researcher working in SYNTHEX project, at System Development Corporation, a computer software company based in Santa Monica, California, founded in 1955 and considered the first company of its kind.

2.3 Ontology in Linguistics

From the linguistics point of view, Schalley (2015) discusses how the world's languages are approached through ontologies. The "prominent notions of 'ontology'" (Schalley, 2015, p.2) encompass the meaning of the term surveyed in philosophy, artificial intelligence, knowledge representation, and, finally, linguistics.

- From the philosophical point of view, Schalley cites Hofweber (2018) to explain that the interest falls under what kinds of things there are in reality;
- From the artificial intelligence point of view, she cites Gruber (1993) to conclude that the object of study in ontologies is not objective but human conceptualization;
- Within the knowledge representation field, the author cited is Sowa (2003), in which language is considered one influencing factor in ontologies. Indeed, Sowa is an author strongly influenced by a semiotic view of ontologies (see, for example, Sowa 2005);
- The point of view of linguistics, which is the genuine interest of Schalley (2015), takes a definition by Schalley, Musgrave, & Haugh (2014) to conclude that the study of ontology refers to human conceptualization, not reality.

If ontology takes either reality or human conceptualization is a decision that should not be directly assigned to the knowledge representation and artificial intelligence fields but on the philosophical theories underlying them. Theories of categories that consider human conceptualization (or concepts) are, in general, gathered under the already mentioned label "conceptualism" (McLear, 2010; Owen, 1991), and theories that consider things in reality use the also already cited label "realism" (MacLeod & Rubenstein, 2005; Niiniluoto, 1999). Thus, one can research either knowledge representation or artificial intelligence and still advocate the study of reality – the things, not their names – in ontologies. Indeed, an extensive and influential chain in ontological studies falls under realism (Grenon & Smith, 2004). There are also relevant long-term initiatives, such as Gene Ontology (Ashburner et al., 2000; Gene Ontology Consortium, 2023).

Schalley (2015) addresses ontologies through a conceptualist view of linguistics. On the other hand, considering the realist view, language is not the focus. As cited at the beginning of a traditional book of ontology as a discipline: "This book is a book about the world. I am concerned with ontology, not merely with language" (Johansson, 1989, p.1). One might say that language is a thing, and then it can be studied through a realist lens. However, language's ontological status is not easy to find and is still debatable (see Scholz 2024). Also, there are other open discussions on themes fundamental for applied ontology without a consensus between philosophy and linguistics, for example, the transitivity of the part-whole relationships (Almeida & Baracho, 2014).

What would be linguistic ontology in the sense of ontology as a discipline? An ontology of this sort could be very useful in harmonizing the myriad notions, concepts and theories within Linguistics. Indeed, there are studies in this sense, for example, in phonology (Odden, 2021). Similar studies come from philosophy and applied ontology, both of which were put forward by philosophers.

Scholz et al. (2024), for example, suggest that the debate on the ontology of language has seen three significant options in the philosophy literature: i) the mentalism of Chomskyan linguistics; ii) a Platonistic (realist) alternative (Katz, 1981), and ii) a nominalist alternative (Devitt 2006).

Dowland et al. (2023) approach some required definitions for language in the context of healthcare systems. In this case, such definitions must fit in a top-level ontology, namely, BFO. Languages would be "aggregates of linguistic competencies," where linguistic competence is a disposition. A disposition is a realizable entity so that each linguistic competence is realized in

certain activities – reading, writing, speaking, etc. – which, in turn, are examples of processes that realize a specific capacity.

To precisely define ontology within linguistics is beyond the goals of this article. Thus, we follow Schalley (2015, p.8) "to illustrate how linguistics might draw on available ontology tools from computer science and build explicit models." To accomplish this goal, we will present an ongoing initiative to build an ontology for corpus linguistics (Section 3).

3 Towards an ontology for corpus linguistics

The C-ORAL-BRASIL Ontology was built as part of a long-term project in linguistic, namely, the C-ORAL-BRASIL project (Raso & Mello, 2012). C-ORAL-BRASIL is a Brazilian Portuguese spontaneous speech corpus, representative of the *Mineiro* (*Minas Gerais* state) diatopy from the metropolitan area of the state capital (*Belo Horizonte* city) (Sardinha & Ferreira, 2014).

We adopted the tenets Arp and Smith (2015) recommended as a methodology regarding domain ontology design, term definitions, classification, top-level entities and relationships. For implementation, we employed the well-known editor Protégé (Musen, 2015) to build the ontology in the Web Ontology Language (OWL) (W3C OWL Working Group, 2012). Linguists of the C-ORAL-BRASIL project provided expertise, which was elicited through discussion with the researchers.

Regarding design, the domain classes are underclasses of the middle—and top-level ontologies that work as primitives. Section 3.1 explains the basic concepts of such ontologies. Section 3.3 presents an excerpt of the main classes of the resulting domain ontology, the C-ORAL-BRASIL Ontology.

3.1 Top and middle-level ontologies employed

In the late 1990s, information technology produced knowledge representation languages whose high capacity for queries and automatic inferences enabled them to operationalize ontologies as artifacts and meet demands that databases could no longer meet. Thus, the first approaches to applied ontology emerged in a context where well-founded design principles and new logical representation languages converged. In addition to modeling languages, such as the Web Ontology Language (OWL), and design principles, such as the principles of the Open Biological and Biomedical Ontology (OBO) Foundry (Smith et al., 2007; Jackson et al., 2021), top-level ontologies have been developed.

Top-level ontologies are ontologies as artifacts, sometimes called upper ontologies, seek to distinguish the most basic categories existing in an ontology (Jansen, 2008). Top-level ontologies are crucial in integrating entities or concepts from different heterogeneous ontologies to represent one domain or use case (Mascardi, Cordì & Rosso 2007). One famous top-level ontology is Basic Formal Ontology (BFO), inspired by Aristotle and Husserl (Grenon & Smith, 2004). We used BFO and middle-level ontologies, extensions of BFO, as the foundation for the CORAL ontology.

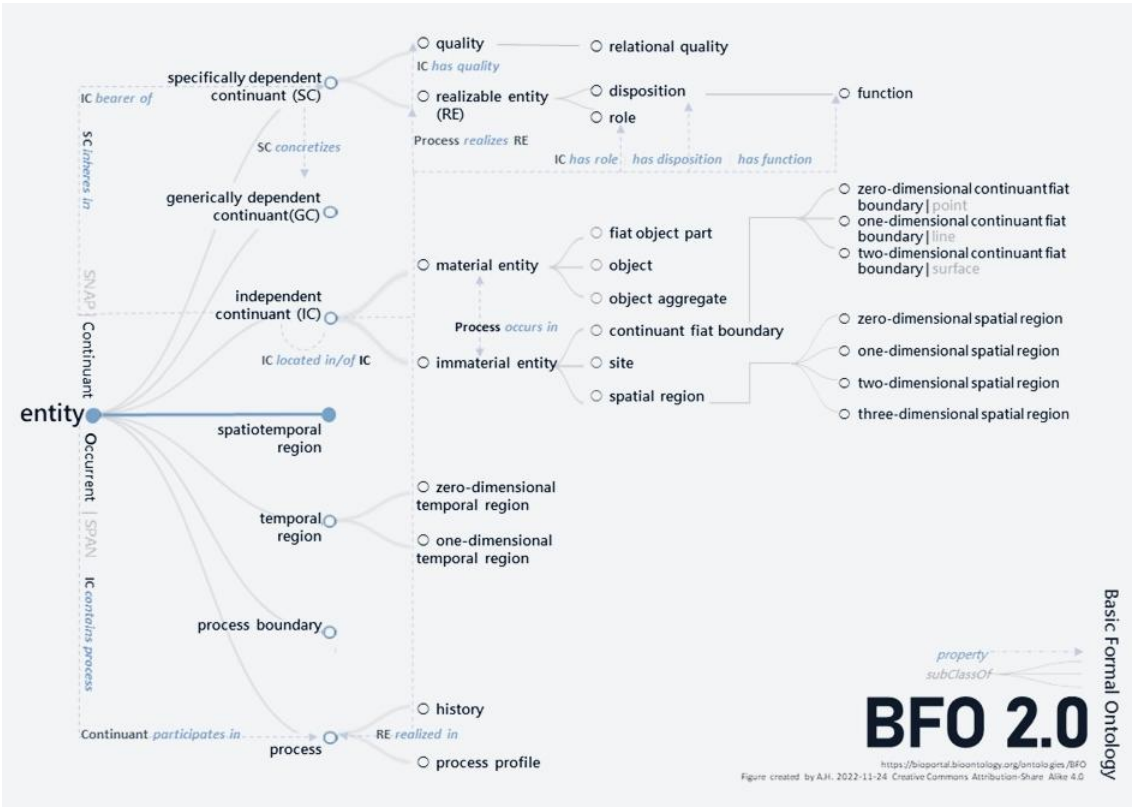

Figure 2: Mind map of BFO, from https://bioportal.bioontology.org/ontologies/BFO

BFO provides a metaphysically grounded perspective on the most basic categories (Figure 2). Through its taxonomy and the provided elucidation for each category, BFO functions as an essential guide for developing ontologies in different knowledge domains. BFO integrates two distinct aspects of existence, categorizing entities into the two taxonomic branches. Specifically, it integrates ontologies that classify *continuants* with those that classify *occurrents* (Grenon, Smith, & Goldberg, 2007) (see Figure 2). A continuant is defined by all its parts being present at each moment that it exists as a whole, whereas the occurrent does have parts that cannot exist synchronously. For example, you are a continuant, but your life is an occurrent.[3]

Scanning the BFO taxonomy makes it possible to find entities with objects of the domain of interest that match while identifying their ontological relationships. However, the match between generic and domain entities can be complex. Therefore, in constructing domain ontologies, ontologists often employ middle-level ontologies. The BFO has several extensions that function as middle-level ontologies. This research used a BFO extension, Information Artifact Ontology (Smith & Ceusters, 2015).

As the name suggests, Information Artifact Ontology (IAO) addresses information artifacts, and its central entity is called "Informational Content Entity" (ICE). One of the primary distinctions underlying BFO's categorization of reality is between independent and dependent continuants. Independent continuants are continuants that do not ontologically depend on another entity.

Examples of ontological dependence include the red of a tomato; for example, the quality of *red* (an instance of that color) depends on the tomato (an instance of an independent continuant). In the case of color, which is a quality, it specifically depends on the tomato. The

[3] Although a full explanation of BFO entities lies beyond the goals of this article, interested readers will find the thorough account in Arp and Smith (2015) and a synthesis in Almeida (2013).

individual color of this tomato ontologically depends on the existence of this tomato. If this tomato goes away, so does its color.

In considering information content entities, we add a new dimension of dependency, namely "generically dependence." This generic dependency means that, for example, an information content entity can inhere in more than one bearer because it is not dependent on any specific bearer, like the color of a specific tomato. On one hand, the "color of a specific tomato" is a particular of the universal "red". Thus, the specific dependency occurs because that instance of red can not migrate to another tomato. On the other hand, the information content of a legal document is not specifically dependent on the original paper document. It can be copied to other paper documents; it can also be transformed into a PDF file or even an audio file. If the original perishes, the information content entity persists, depending on the other forms of carriers. We call entities that, like Information Content Entities (ICEs), are ontologically dependent but not on any specific bearer by the label generically dependent continuants (GDCs).

Domain terms, namely, the terms related to corpus linguistics, are under the top-level and mid-level ontologies. For example, "tag" is a central term within the domain since it subsumes all marks used to identify linguistic sentences and expressions of interest in the corpus. So, following the full path, from "entity" to "tag", one will find:

1. From the top-level: entity > continuant > generically dependent continuant...
2. From the middle-level: ... information content entity > textual entity...
3. From the domain-level: ... tag ...

All this machinery explained so far is employed to model the generic part that grounds the domain ontology for corpus linguistics.

3.2 Classes and Relationships in the C-ORAL-BRASIL Ontology

The rationale for developing the domain ontology – C-ORAL-BRASIL Ontology – follows the expertise of linguistics researchers who provided specialized knowledge. The domain of interest starts with a speaker, an agent who participates in an utterance process. Such a speaker has the role of participant that is realized by the utterance process. The participant can be formal or informal. The utterance process has an output: the utterance energy. The utterance energy is a material entity since it is sound energy. The utterance process still realizes some speech and has an utterance as a participant. The utterance is a kind of information artifact entity; the document part and textual entity are information artifact entities, too. Metadata is a kind of textual entity: metadata of the first order is about something in reality; metadata of the second order is about other metadata. The corpus is about the kinds of speech and is composed of texts. Texts are document parts and can be tagged text and non-tagged texts. This kind of tag delimits linguistic phenomena of interest for linguists and researchers.

In the remainder of this section, we describe some of the main classes created in the CORAL-Ontology since it is impossible to depict the whole ontology (Figure 2). The ontology is documented on the Internet[4], and more information about the classes' definitions and relationships presented here is available in the Appendix.

[4] Available on the Internet: https://mba.eci.ufmg.br/coral/doc/index-en.html

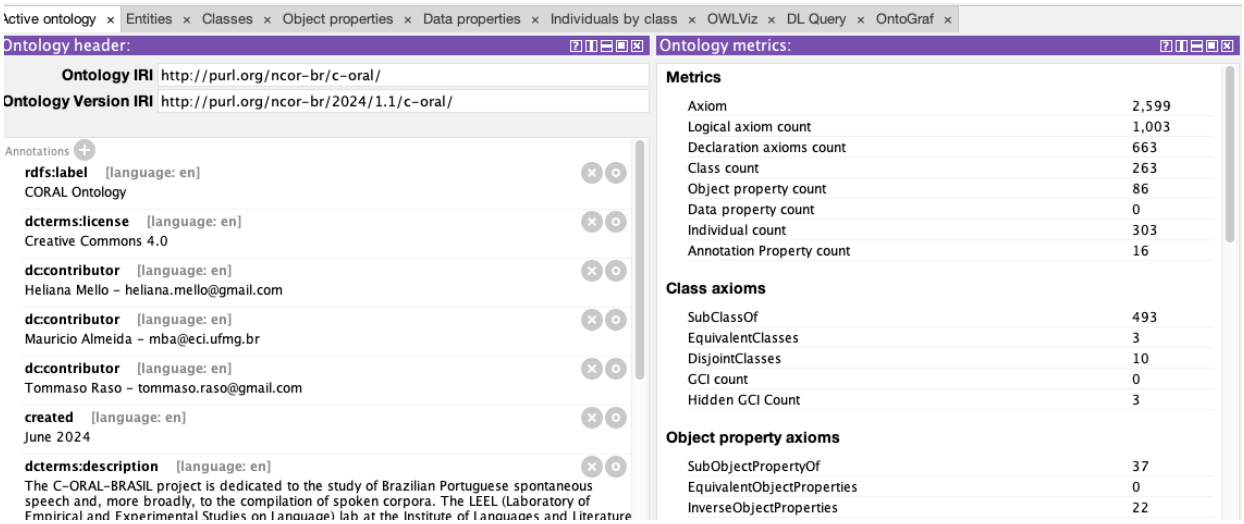

Figure 2: Metrics for the CORAL-Ontology with 263 classes and 86 relationships

To understand the path in the ontology to reach the "tag" class – an essential class for the corpus – we can start with the "agent" class following the taxonomic (is-a) and other relationships (Figure 3a):

- A speaker is an agent in the utterance process.
- An utterance process is part of a speech process.
- An utterance process has participants an utterance.
- An utterance is an information-content entity.
- A textual entity is an information context entity.
- A tag is a textual entity.
- All kinds of tags are defined under the tag class (Figure 3b)

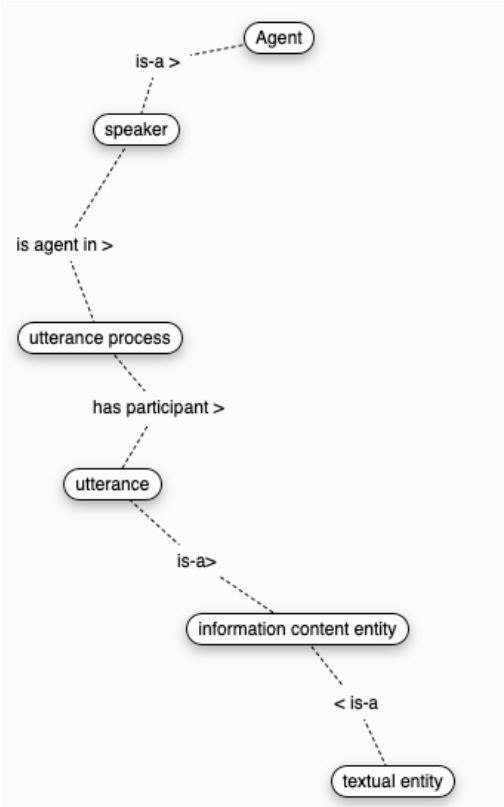

Figure 3a: From agent class to textual entity class

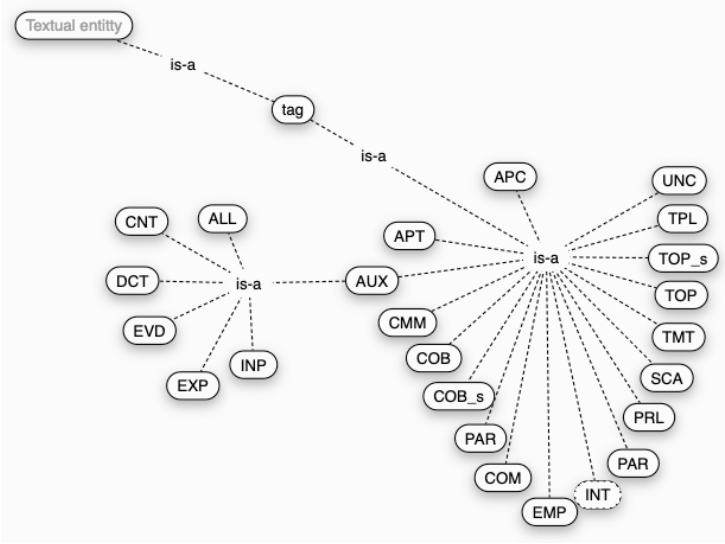

Figure 3b: From textual entity to tag class

Another relevant group of entities is included in the "informational unit" class. Tags mark informational units; for example, an allocutive is marked by the ALL tag, and an appendix of comment is marked by an APC tag (Figure 4, see definitions in the Appendix):

- We can find all informational unit classes under the prosodic unit class.
- Dialogic auxiliary and textual unit classes are subclasses of the informational unit class.
- The dialogic auxiliary has the following classes as subclasses: allocutive, conative, evidentiator, expressive, and incipit.
- The textual unit has the following classes as subclasses: appendix of comment, appendix of topic, discourse connector, illocutionary unit, list of parenthetic, locutive introducer,

parenthetic subordinator bound comment, topic, topic list, topic subordinator, bound comment, comment and multiple comments.

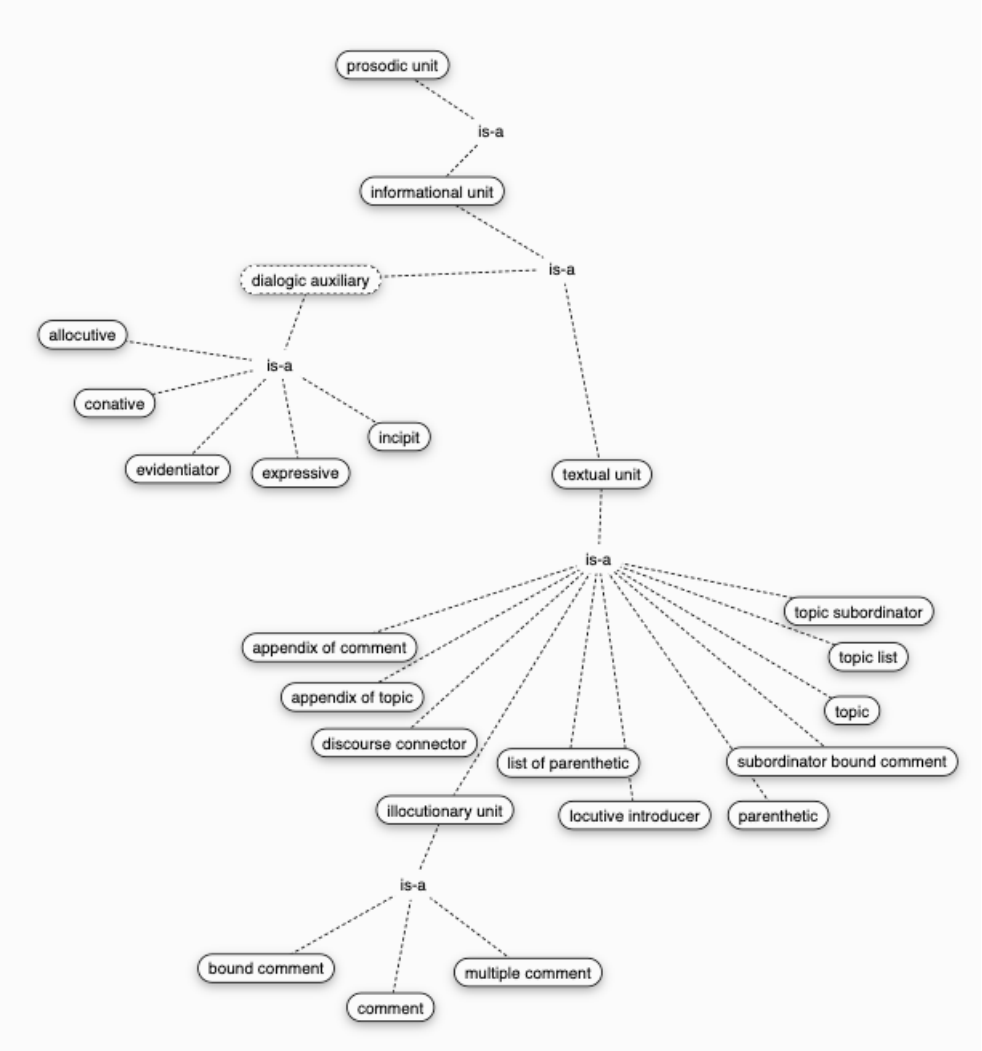

Figure 4: Classes under informational unit

The "utterance process" and "textual entity" are other essential parts of the ontology. It starts with a speaker, an agent in an utterance process (Figure 5). Definitions can be found in the Appendix:

- An utterance process is part of the speech process.
- The speech process is recorded in a corpus.
- A corpus is composed of some corpus text.
- Corpus text can be tagged text or nontagged text.
- A corpus text is a textual entity.
- The textual entity is an information content entity.
- The textual entity class has as subclasses metadata, segment of text, tag, and word-level sequence.

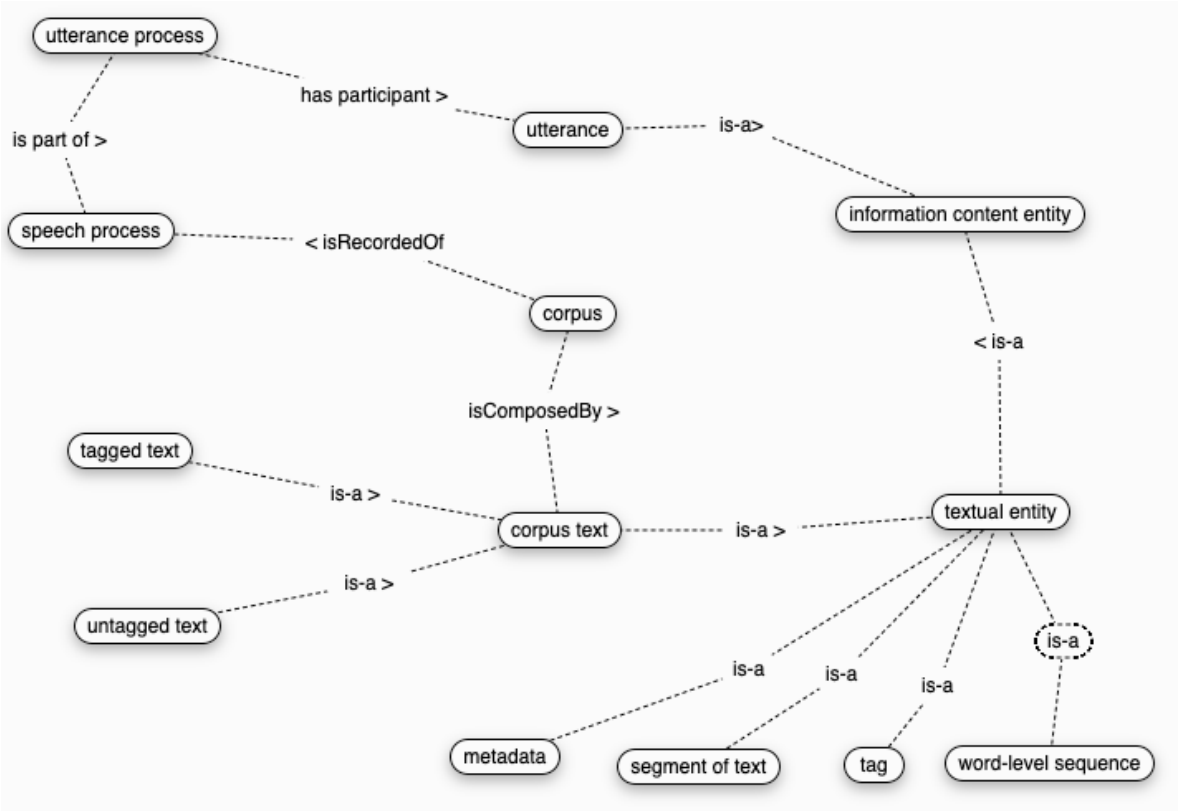

Figure 6: Classes related to the "Utterance process" class

The C-ORAL-BRASIL project also has a branch to interconnect speech and diseases like schizophrenia. The goal is to investigate how schizophrenia can impact the speech of people. Important notions in this context are psychopathology, description of symptoms (positive, negative, cognition), and diagnostic criteria. The Open Biological and Biomedical Ontology (OBO) Foundry (Smith et al., 2007), an already mentioned suite of biomedical ontologies, has an ontology specifically built for schizophrenia called Mental Disease Ontology (MDO)[5]. Part of MDO (Figure 6) is reused within C-ORAL-BRASIL Ontology.

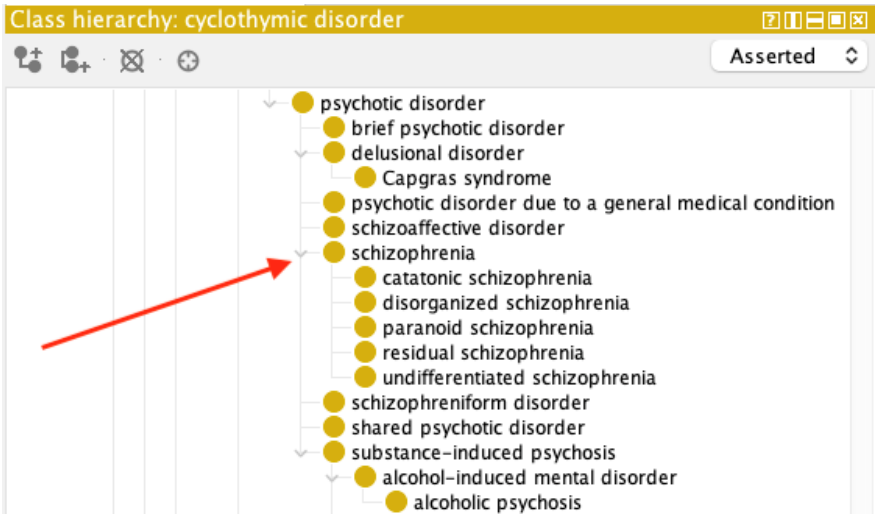

Figure 7: Fragment of MDO highlighting schizophrenia class and its subclasses

[5] Available on the Internet: https://obofoundry.org/ontology/mfomd.html

4 Final remarks

In this article, we began introducing the interdisciplinary question: What is an ontology? We explained some views from philosophy, computer science and linguistics. We also explained the top and middle levels that ground the CORAL ontology, a corpus linguistics ontology under construction. Then, we presented definitions, labels and relationships of the resulting ontology counts with 263 classes and 86 relationships distributed over three layers: domain, middle, and top-level ontologies.

Studies that aim to understand the nature of language are very relevant topics. The present paper follows this direction by approaching language through an applied ontological approach. The top-level ontologies introduced use principles coming from the ontology as a discipline; the C-ORAL-BRASIL Ontology was built according to methods to create ontologies as artifacts.

In future works, we hope to clarify the uses of ontology as an artifact reported in the article in the context of information systems. Thus, we may show the direct utility in building ontologies for linguists, computer and information scientists and even ontologists.

Acknowledgments

This work was supported by the *Instituto de Estudos Avançados Transdisciplinares* (IEAT), *Universidade Federal de Minas Gerais* (UFMG), Brazil.

REFERENCES

1. Almeida M. Revisiting Ontologies: A Necessary Clarification, Journal of the American Society for Information Science and Technology [Internet]. 2013 May [cited 2024 Dec 12]; 64(8): [about 11pp.]. Available from: https://onlinelibrary.wiley.com/doi/full/10.1002/asi.22861

2. Almeida M B, Baracho, R M A. A Theoretical Investigation About The Notion Of Parts And Wholes: Mereological And Meronymic Relations. Brazilian Journal of Information Science: Research Trends [Internet]. 2014 May [cited 2024 Dec 12]; 8(1/2): [about 40pp.]. Available from: https://revistas.marilia.unesp.br/index.php/bjis/article/view/4247

3. Almeida M. *Ontology in Information Science*: Theory and Method. Curitiba: CRV, 2020 (In Portuguese).

4. Arp R, Smith B, Spear AD. *Building Ontologies with Basic Formal Ontology*. Cambridge: The MIT Press, 2015.

5. Ashburner M, et al. Gene ontology: tool for the unification of biology. Nature Genetics [Internet]. 2000 May [cited 2024 Dec 12]; 25(9): [about 4pp.]. Available from: https://www.nature.com/articles/ng0500_25

6. Beyer C. Edmund Husserl. In: The Stanford Encyclopedia of Philosophy [Internet]. Metaphysics Research Lab, Stanford University (CA). [last modified 2017; cited 2024 Dec 05]. Available from: http://plato.stanford. edu/entries/husserl/

7. Devitt M. *Ignorance of Language*. Oxford: Clarendon Press, 2006.

8. Hofweber T. Logic and ontology. In: The Stanford Encyclopedia of Philosophy [Internet]. Metaphysics Research Lab, Stanford University (CA). [last modified 2017; cited 2024 Dec 05]. Available from https://plato.stanford.edu/archives/sum2018/entries/logic-ontology/

9. The Gene Ontology Consortium. The Gene Ontology knowledgebase in 2023. Genetics [Internet]. 2023 May [cited 2024 Dec 12]; 224(1): [about pp.]. Available from: https://pubmed.ncbi.nlm.nih.gov/36866529/

10. Grenon P, Smith B. SNAP and SPAN: Towards Dynamic Spatial Ontology. Spatial Cognition & Computation, [Internet]. 2009 Nov [cited 2024 Dec 12]; 4(1): [about 35pp.]. Available from: https://www.tandfonline.com/doi/abs/10.1207/s15427633scc0401_5

11. Gruber T R. A translation approach to portable ontology specifications. Knowledge Acquisition, [Internet]. 1993 May [cited 2024 Dec 12]; 5(2): [about 21pp.]. Available from: https://www.sciencedirect.com/science/article/pii/S1042814383710083

12. Guarino N. Formal ontology in information systems. [Internet]. [Place unknown]; 1998 [cited 2024 Dec 15]. Available from: https://www.loa.istc.cnr.it/old/Papers/FOIS98.pdf

13. Guizzardi G. Ontological foundations for structural conceptual models. Enschede: Telematica Instituut, 2005.

14. Hennig B. What is formal ontology? In: Munn, K. Smith, B. *Applied ontology: An Introduction*. Berlin: Verlag, 2008. p. 39–56.

15. Jansen L. Categories: The top level ontology. In: Munn, K. Smith, B. *Applied ontology: An introduction*. Berlin: Verlag, 2008. p. 173–196.

16. Johansson I. *Ontological investigations*: an inquiry into the categories of nature, man, and society. New York: Routledge, 1989.

17. Katz J J. Language and Other Abstract Objects, Totowa: Rowman and Littlefield, 1981.

18. McLear C. Kant: Philosophy of Mind. 2010. In: Internet Encyclopedia of Philosophy [Internet]. The Internet Encyclopedia of Philosophy. [last modified 2010; cited 2024 Dec 05]. Disponível em: https://www.iep.utm.edu/kandmind/.

19. MacLeod M, Rubenstein E. Universals. In: Internet Encyclopedia of Philosophy [Internet]. The Internet Encyclopedia of Philosophy (IEP). [last modified 2005; cited 2024 Dec 05]. Disponível em: http://www.iep.utm.edu/universa/

20. Moltmann F. Natural Language Ontology In: The Stanford Encyclopedia of Philosophy [Internet]. Metaphysics Research Lab, Stanford University (CA). [last modified 2022 Aug 17; cited 2024 Dec 05]. Available from https://plato.stanford.edu/archives/spr2024/entries/natural-language-ontology/

21. Musen M A. The Protégé project: A look back and a look forward. AI Matters [Internet]. 2015 Jun [cited 2024 Dec 12]; 1(4): [about 8 pp.]. Available from: https://pubmed.ncbi.nlm.nih.gov/27239556/

22. Niiniluoto I. *Critical scientific realism*. Oxford: Oxford Scholarship, 1999.

23. Owen D. Locke on real essence. History of Philosophy Quarterly [Internet]. 1991 Jun [cited 2024 Dec 12]; 8: [about 13 pp.]. Available from: https://www.jstor.org/stable/27743968

24. Musen M A. The Protégé project: A look back and a look forward. AI Matters [Internet]. 2015 Jun [cited 2024 Dec 12]; 1(4): [about 8 pp.]. Available from: https://pubmed.ncbi.nlm.nih.gov/27239556/

25. Odden D. Phonological Ontology. [Internet]. [Place unknown]; 2021 [updated 2012; cited 2024 Dec 15]. Available from: https://zenodo.org/records/5514165

26. Raso T, Mello H. *C-ORAL-BRASIL I: Corpus de Referência do Português Brasileiro Falado Informal*. Belo Horizonte: UFMG, 2012.

27. Rodriguez-Pereyra G. Nominalism in Metaphysics. In: The Stanford Encyclopedia of Philosophy [Internet]. Metaphysics Research Lab, Stanford University (CA). [last modified 2015 Apr; cited 2024 Dec 15]. Available from: https://plato.stanford.edu/archives/spr2024/entries/nominalism-metaphysics/

28. Sardinha T B, Ferreira T L S B. *Working with Portuguese Corpora*. London: Bloomsbury, 2014.

29. Schalley AC. Ontologies and ontological methods in linguistics. Language Linguist Compass. [Internet]. 2019 Aug; cited 2024 Dec 15]; 13(11): [about 19 pp.]. Available from: https://compass.onlinelibrary.wiley.com/doi/10.1111/lnc3.12356

30. Schalley A C, Musgrave S, Haugh M. Accessing phonetic variation in spoken language corpora through non-standard orthography. Australian Journal of Linguistics, [Internet]. 2014 Feb [cited 2024 Dec 15]; 34(1): [about 31 pp.]. Available from: https://espace.library.uq.edu.au/view/UQ:388403

31. Sowa J. Conceptual Graphs. [Internet]. [Place unknown]: 2005 [updated 2005 Mar 14; cited 2024 Dec 12]. Available from: https://www.jfsowa.com/cg/index.htm

32. Scholz BC, Pelletier FJ, Pullum, GK, Nefdt, PR. Philosophy of Linguistics, In: The Stanford Encyclopedia of Philosophy [Internet]. Metaphysics Research Lab, Stanford University (CA). [last modified 2024; cited 2024 Dec 05]. Available from https://plato.stanford.edu/archives/spr2024/entries/linguistics/.

33. Smith B. Ontology and information systems. [Internet]. [Place unknown]: 2004 [updated 2004; cited 2024 Dec 12]. Available from: https://ontology.buffalo.edu/ontology(PIC).pdf

34. Smith B. et al. The OBO Foundry: coordinated evolution of ontologies to support biomedical data integration. Nature Biotechnology [Internet]. 2007 Jun [cited 2024 Dec 12]; 25(11): [about 4 pp.]. Available from: https://www.nature.com/articles/nbt1346

35. Staab S, Studer R. *Handbook on ontologies*. Berlin: Springer, 2004.

36. Stefik M. *Introduction to knowledge systems*. San Francisco: Morgan Kaufmann, 1995.

37. Thomason A. Categories. In: The Stanford Encyclopedia of Philosophy [Internet]. Metaphysics Research Lab, Stanford University (CA). [last modified 2022 Nov; cited 2024 Dec 15]. Available from: http://plato.stanford.edu/entries/categories/

38. Wand Y, Storey V C, Weber R. An ontological analysis of the relationship construct in conceptual modeling. ACM Transactions on Database Systems, [Internet]. 1999 Dec [cited 2024 Dec 12]; 24(4): [about 32 pp.]. Available from: https://dl.acm.org/doi/10.1145/331983.331989

39. Wand Y, Weber R. Mario Bunge's ontology as a formal foundation for information systems concepts. In: Studies on Mario Bunge's Treatise . Amsterdam: Rodopi, 1990. p. 123–149.

40. W3Consortium OWL Working Group. [Internet]. [Place unknown]: 2012 [updated 2012 Dec 11; cited 2024 Dec 12]. Available from: https://www.w3.org/TR/owl2-overview/

Appendix

Class = label of the class

Definition = definition of the class in natural language

Relationship = axioms in OWL (Manchester Syntax)[6]

Classes under "Tag" (See Figure 2).

Class	ALL
Definition	ALL is a AUX tag that has function of marking an allocutive in the illocutionary level.
Relationship	(marks some allocutive) and (pertain some 'segment of text illocutionary level')

Class	ALL-PAR
Definition	ALL-PAR is a AUX-PAR tag that has function of marking an allocutive in the parenthetical level.
Relationship	(marks some allocutive) and (pertain some 'segment of text parenthetical level')

Class	APC
Definition	APC is a tag that has function of marking an appendix of comment in the illocutionary level.
Relationship	(marks some 'appendix of comment') and (pertain some 'segment of text illocutionary level')

Class	APC-PAR
Definition	APC is a tag that has function of marking an appendix of comment in the parenthetical level.
Relationship	(marks some 'appendix of comment') and (pertain some 'segment of text parenthetical level')

Class	APC-r
Definition	APC is a tag that has function of marking an appendix of comment in the metaillocutionary level.
Relationship	(marks some 'appendix of comment') and (pertain some 'segment of text metaillocutionary level')

Class	APT
Definition	APT is a tag that has function of marking an appendix of topic in the parenthetical level.
Relationship	(marks some 'appendix of topic') and (pertain some 'segment of text illocutionary level')

Class	APT-r
Definition	APT is a tag that has function of marking an appendix of topic in the metaillocutionary level.

[6] Avaialable on the internet: https://www.w3.org/TR/owl2-manchester-syntax/

Relationship	(marks some 'appendix of topic') and (pertain some 'segment of text metaillocutionary level')

Class	Aux
Definition	AUX is a tag that has function of marking a dialogic auxiliary in the illocutionary level.
Relationship	(marks some 'dialogic auxiliary') and (pertain some 'segment of text illocutionary level')

Class	AUX-PAR
Definition	AUX-PAR is a tag that has function of marking a dialogic auxiliary in the parenthetical level.
Relationship	(marks some 'dialogic auxiliary') and (pertain some 'segment of text parenthetical level')

Class	AUX-r
Definition	AUX-r is a tag that has function of marking a dialogic auxiliary in the metaillocutionary level.
Relationship	(marks some 'dialogic auxiliary') and (pertain some 'segment of text metaillocutionary level')

Class	CMM
Definition	CMM is a tag that has function of marking a multiple comment in the illocutionary level.
Relationship	(marks some 'multiple comment') and (pertain some 'segment of text illocutionary level')

Class	CMM-PAR
Definition	CMM-PAR is a tag that has function of marking a multiple comment in the parenthetical level.
Relationship	(marks some 'multiple comment') and (pertain some 'segment of text parenthetical level')

Class	CMM-r
Definition	CMM-r is a tag that has function of marking a multiple comment in the metaillocutionary level.
Relationship	(marks some 'multiple comment') and (pertain some 'segment of text metaillocutionary level')

Class	CNT
Definition	CNT is a AUX tag that has function of marking a conative in the illocutionary level.
Relationship	(marks some conative) and (pertain some 'segment of text illocutionary level')

Class	CNT-PAR
Definition	CNT-PAR is a AUX-PAR tag that has function of marking a conative in the parenthetical level.
Relationship	(marks some conative) and (pertain some 'segment of text parenthetical level')

Class	CNT-r
Definition	CNT-r is a AUX-r tag that has function of marking a conative in the metaillocutionary level.
Relationship	(marks some conative) and (pertain some 'segment of text metaillocutionary level')

Class	COB
Definition	COB is a tag that has function of marking a bound comment in the illocutionary level.
Relationship	(marks some 'bound comment') and (pertain some 'segment of text illocutionary level')

Class	COB-PAR
Definition	COB-PAR is a tag that has function of marking a bound comment in the parenthetical level.
Relationship	(marks some 'bound comment') and (pertain some 'segment of text parenthetical level')

Class	COB-r
Definition	COB-r is a tag that has function of marking a bound comment in the metaillocutionary level.
Relationship	(marks some 'bound comment') and (pertain some 'segment of text metaillocutionary level')

Class	COB_s
Definition	COB is a tag that has function of marking a subordinator bound comment in the illocutionary level.
Relationship	(marks some 'subordinator bound comment') and (pertain some 'segment of text illocutionary level')

Class	COB_s-PAR
Definition	COB_s-PAR is a tag that has function of marking a subordinator bound comment in the parenthetical level.
Relationship	(marks some 'subordinator bound comment') and (pertain some 'segment of text parenthetical level')

Class	COB_s-r
Definition	COB_s-r is a tag that has function of marking a subordinator bound comment in the metaillocutionary level.
Relationship	(marks some 'subordinator bound comment') and (pertain some 'segment of text metaillocutionary level')

Class	COM
Definition	COM is a tag that has function of marking a comment in the illocutionary level.
Relationship	(marks some comment) and (pertain some 'segment of text illocutionary level')

Class	COM-PAR

Definition	COM-PAR is a tag that has function of marking a comment in the parenthetical level.
Relationship	(marks some comment) and (pertain some 'segment of text parenthetical level')

Class	COM-r
Definition	COM-r is a tag that has function of marking a comment in the metaillocutionary level.
Relationship	(marks some comment) and (pertain some 'segment of text metaillocutionary level')

Class	DCT
Definition	DCT is a tag that has function of marking a discourse connector in the illocutionary level.
Relationship	(marks some 'discourse connector') and (pertain some 'segment of text illocutionary level')

Class	DCT-PAR
Definition	DCT-PAR is a tag that has function of marking a discourse connector in the parenthetical level.
Relationship	(marks some 'discourse connector') and (pertain some 'segment of text parenthetical level')

Class	DCT-r
Definition	DCT-r is a tag that has function of marking a discourse connector in the metaillocutionary level.
Relationship	

Class	EMP
Definition	EMP is a tag that has function of marking an empty.
Relationship	(marks some 'discourse connector') and (pertain some 'segment of text metaillocutionary level')

Class	EVD
Definition	EVD is a AUX tag that has function of marking an evidentiator in the illocutionary level.
Relationship	(marks some evidentiator) and (pertain some 'segment of text illocutionary level')

Class	EVD-PAR
Definition	EVD-PAR is a AUX-PAR tag that has function of marking an evidentiator in the parenthetical level.
Relationship	(marks some evidentiator) and (pertain some 'segment of text parenthetical level')

Class	EVD-r
Definition	EVD-r is a AUX-r tag that has function of marking an evidentiator in the metaillocutionary level.
Relationship	(marks some evidentiator) and (pertain some 'segment of text metaillocutionary level')

Class	EXP
Definition	EXP is a AUX tag that has function of marking an expressive in the illocutionary level.
Relationship	(marks some expressive) and (pertain some 'segment of text illocutionary level')

Class	EXP-PAR
Definition	EXP-PAR is a AUX-PAR tag that has function of marking an expressive in the parenthetical level.
Relationship	(marks some expressive) and (pertain some 'segment of text parenthetical level')

Class	EXP-r
Definition	EXP-r is a AUX-r tag that has function of marking an expressive in the metaillocutionary level.
Relationship	(marks some expressive) and (pertain some 'segment of text metaillocutionary level')

Class	INP
Definition	INP is a AUX tag that has function of marking an incipit in the illocutionary level.
Relationship	(marks some incipit) and (pertain some 'segment of text illocutionary level')

Class	INP-PAR
Definition	INP-PAR is a AUX-PAR tag that has function of marking an incipit in the parenthetical level.
Relationship	(marks some incipit) and (pertain some 'segment of text parenthetical level')

Class	INP-r
Definition	INP-r is a AU-r tag that has function of marking an incipit in the metaillocutionary level.
Relationship	(marks some incipit) and (pertain some 'segment of text metaillocutionary level')

Class	INT
Definition	INT is a tag that has function of marking a locutive introducer in the illocutionary level.
Relationship	(marks some 'locutive introducer') and (pertain some 'segment of text illocutionary level')

Class	INT-PAR
Definition	INT-PAR is a tag that has function of marking a locutive introducer in the parenthetical level.
Relationship	(marks some 'locutive introducer') and (pertain some 'segment of text parenthetical level')

Class	INT-r

Definition	INT-r is a tag that has function of marking a locutive introducer in the metaillocutionary level.
Relationship	(marks some 'locutive introducer') and (pertain some 'segment of text metaillocutionary level')

Class	PAR
Definition	PAR is a tag that has function of marking a parenthetic in the illocutionary level.
Relationship	(marks some parenthetic) and (pertain some 'segment of text illocutionary level')

Class	PAR-PAR
Definition	PAR-PAR is a tag that has function of marking a parenthetic in the parenthetical level.
Relationship	

Class	PAR-r
Definition	PAR-r is a tag that has function of marking a parenthetic in the metaillocutionary level.
Relationship	(marks some parenthetic) and (pertain some 'segment of text parenthetical level')

Class	PRL
Definition	PRL is a tag that has function of marking a list of parenthetic in the illocutionary level.
Relationship	(marks some 'list of parenthetic') and (pertain some 'segment of text illocutionary level')

Class	PRL-PAR
Definition	PRL-PAR is a tag that has function of marking a list of parenthetic in the parenthetical level.
Relationship	(marks some 'list of parenthetic') and (pertain some 'segment of text parenthetical level')

Class	PRL-r
Definition	PRL-r is a tag that has function of marking a list of parenthetic in the metaillocutionary level.
Relationship	(marks some 'list of parenthetic') and (pertain some 'segment of text metaillocutionary level')

Class	SCA
Definition	SCA is a tag that has function of marking a scanning in the illocutionary level.
Relationship	(marks some scanning) and (pertain some 'segment of text illocutionary level')

Class	SCA-PAR
Definition	SCA-PAR is a tag that has function of marking a scanning in the parenthetical level.

Relationship	(marks some scanning) and (pertain some 'segment of text parenthetical level')

Class	SCA-r
Definition	SCA-r is a tag that has function of marking a scanning in the metaillocutionary level.
Relationship	(marks some scanning) and (pertain some 'segment of text metaillocutionary level')

Class	TMT
Definition	TMT is a tag that has function of marking a time taking.
Relationship	(marks some 'time taking') and (pertain some 'segment of text illocutionary level')

Class	TOP
Definition	TOP is a tag that has function of marking a topic in the illocutionary level.
Relationship	(marks some 'time taking') and (pertain some 'segment of text illocutionary level')

Class	TOP-PAR
Definition	TOP-PAR is a tag that has function of marking a topic in the parenthetical level.
Relationship	(marks some topic) and (pertain some 'segment of text parenthetical level')

Class	TOP-r
Definition	TOP-r is a tag that has function of marking a topic in the metaillocutionary level.
Relationship	(marks some topic) and (pertain some 'segment of text metaillocutionary level')

Class	TOP_s
Definition	TOP_s is a tag that has function of marking a topic subordinator.
Relationship	(marks some topic) and (pertain some 'segment of text parenthetical level')

Class	TPL
Definition	TPL is a tag that has function of marking a topic list in the illocutionary level.
Relationship	(marks some 'topic list') and (pertain some 'segment of text illocutionary level')

Class	TPL-PAR
Definition	TPL-PAR is a tag that has function of marking a topic list in the parenthetical level.
Relationship	(marks some 'topic list') and (pertain some 'segment of text parenthetical level')

Class	TPL-r
Definition	TPL-r is a tag that has function of marking a topic list in the metaillocutionary level.
Relationship	(marks some 'topic list') and (pertain some 'segment of text metaillocutionary level')

Class	UNC
Definition	UNC is a tag that has function of marking an unclear in the illocutionary level.
Relationship	(marks some unclear) and (pertain some 'segment of text illocutionary level')

Class	UNC-PAR
Definition	UNC-PAR is a tag that has function of marking an unclear in the parenthetical level.
Relationship	(marks some unclear) and (pertain some 'segment of text parenthetical level')

Class	UNC-r
Definition	UNC-r is a tag that has function of marking an unclear in the metaillocutionary level.
Relationship	(marks some unclear) and (pertain some 'segment of text metaillocutionary level')

Classes under "Informational Unit" (See Figure 3).

Class	Prosodic unit
Definition	Prosodic unit is a sequence of words in a text followed by /, //, + or [/n] that represent sequence in a speech.
Relationship	-

Class	Informational unit
Definition	Informational unit is prosodic unit that represents an informational value.
Relationship	-

Class	Dialogic auxiliary
Definition	Dialogic auxiliary is an informational unit that has the function to regulate the interaction in an illocutionary, metaillocutionary or parenthetical level.
Relationship	isMarkedBy some (ALL or CNT or EVD or EXP or INP) pertain some ('segment of text illocutionary level' or 'segment of text metaillocutionary level' or 'segment of text parenthetical level')

Class	Allocutive
Definition	Allocutive is a dialogic auxiliary that has the function to mark the social cohesion with respect to the interlocutor.
Relationship	isMarkedBy only ALL

Class	Conative
Definition	Conative is a dialogic auxiliary that that has the function to mark the illocutionary solution.
Relationship	isMarkedBy only CNT

Class	Evidentiator
Definition	Evidentiator is a dialogic auxiliary that has the function to mark the semantics of previous words.
Relationship	isMarkedBy only EVD

Class	Expressive
Definition	Expressive is a dialogic auxiliary that has the function to signal surprise with respect to previous content.
Relationship	isMarkedBy only EXP

Class	Incipt
Definition	Incipit is a dialogic auxiliary that has the function to signal the beginning of turn or pattern.
Relationship	isMarkedBy some INP

Class	Textual entity
Definition	A textual entity is a part of a manifestation (FRBR sense), a generically dependent continuant whose concretizations are patterns of glyphs intended to be interpreted as words, formulas, etc.
Relationship	pertain some ('segment of text illocutionary level' or 'segment of text metaillocutionary level' or 'segment of text parenthetical level')

Class	Appendix of comment
Definition	Appendix of comment is a textual unit that has the function to integrate the comment unit.
Relationship	isMarkedBy only APC

Class	Appendix of topic
Definition	Appendix of topic is a textual unit that has the function to integrate the topic unit.
Relationship	isMarkedBy only APT

Class	Discourse connector
Definition	Discourse connector is a textual information unit that has the function to connect two different information patterns.
Relationship	isMarkedBy only DCT

Class	Illocutionary unit
Definition	Illocutionary unit is a textual unit that carries the action performed in the pattern.
Relationship	-

Class	List of parenthetic
Definition	List of parenthetic is textual unit that consists a patternized sequence of parenthetics.
Relationship	isMarkedBy only 'list of parenthetic'

Class	Locutive introducer
Definition	Locutive Introducer is a textual unit that has the function to introduce a metaillocution or a list.
Relationship	isMarkedBy only INT

Class	Parenthetic

Definition	Parenthetic is a textual unit that has the function to add commentaries about the content of the pattern.
Relationship	

Class	Subordinator bound comment
Definition	Subordinated bound comment is a textual unit that governs other comments.
Relationship	isMarkedBy only PAR

Class	Topic
Definition	Topic is a textual unit that indicates the field of application of the illocutionary unit.
Relationship	isMarkedBy only TOP

Class	Topic list
Definition	Topic list is a textual unit that consists of a patternized sequence of topics.
Relationship	isMarkedBy only TPL

Class	Topic subordinator
Definition	Topic subordinator is a textual unit that consist of a topic that governs other topics.
Relationship	isMarkedBy only TOP_s

Classes related to the "Utterance process" (See Figure 4).

Class	Utterance process
Definition	Utterance process is a planned process of making speech sounds which may or may not have an actual language involved.
Relationship	'has participant' some utterance hasSpecifiedOutput some 'utterance energy' partOf some 'speech process'

Class	Speech process
Definition	Speech process is a planned process of making speech sounds which have an actual language involved for purpose of communication.
Relationship	hasMetadata some 'textual metadata'

Class	Corpus
Definition	Corpus is an information content entity that consists of a set of texts in electronic format, compiled according to specific criteria and considered representative of what is to be studied.
Relationship	isComposedBy only 'corpus text' isRecordedOf only 'speech process'

Class	Corpus text
Definition	Corpus text is a textual entity that is concretized in glyphs intended to be interpreted as words.
Relationship	hasSource only 'audio file' isComposedBy only 'segment of text'

Class	Tagged text

Definition	Tagged text is a corpus text that is marked with tags
Relationship	isMarkedBy only tag

Class	Non-tagged text
Definition	Non tagged text is a corpus text that is not marked with tags
Relationship	not (isMarkedBy some tag)

Class	Utterance
Definition	An utterance is an information content entity that is a complete unit of speech in spoken language.
Relationship	-

Class	Information content entity
Definition	A generically dependent continuant that is about some thing.
Relationship	isAbout some entity

Class	Textual entity
Definition	A textual entity is a part of a manifestation (FRBR sense), a generically dependent continuant whose concretizations are patterns of glyphs intended to be interpreted as words, formulas, etc.
Relationship	isAbout some entity

Class	Word-level sequence
Definition	Word-level sequence is a textual entity wich is defined by a word.
Relationship	isAbout some entity 'continuant part of' only continuant

Class	Tag
Definition	Tag is a textual entity that is assigned to a piece of text to allow its identification, organization or retrieval.
Relationship	marks only ('corpus text' or 'segment of text')

Class	Segment of text
Definition	A segment of text is a textual entity that can be delineated based on syntactic, semantic, functional, or structural criteria; and represents a portion of a written or spoken discourse.
Relationship	isAbout some entity 'continuant part of' only continuant

Class	Metadata
Definition	Metadata is a textual entity that is about another information content entity.
Relationship	isAbout some entity 'continuant part of' only continuant