frontiers in Psychology

HYPOTHESIS AND THEORY
published: 23 February 2016
doi: 10.3389/fpsyg.2016.00196

Specialized Knowledge Representation and the Parameterization of Context

Pamela Faber and Pilar León-Araúz*

Department of Translation and Interpreting, University of Granada, Granada, Spain

Though instrumental in numerous disciplines, context has no universally accepted definition. In specialized knowledge resources it is timely and necessary to parameterize context with a view to more effectively facilitating knowledge representation, understanding, and acquisition, the main aims of terminological knowledge bases. This entails distinguishing different types of context as well as how they interact with each other. This is not a simple objective to achieve despite the fact that specialized discourse does not have as many contextual variables as those in general language (i.e., figurative meaning, irony, etc.). Even in specialized text, context is an extremely complex concept. In fact, contextual information can be specified in terms of scope or according to the type of information conveyed. It can be a textual excerpt or a whole document; a pragmatic convention or a whole culture; a concrete situation or a prototypical scenario. Although these versions of context are useful for the users of terminological resources, such resources rarely support context modeling. In this paper, we propose a taxonomy of context primarily based on scope (local and global) and further divided into syntactic, semantic, and pragmatic facets. These facets cover the specification of different types of terminological information, such as predicate-argument structure, collocations, semantic relations, term variants, grammatical and lexical cohesion, communicative situations, subject fields, and cultures.

Keywords: context parameters, specialized knowledge, terminology, terminological knowledge bases

OPEN ACCESS

Edited by:
Marco Cruciani,
University of Trento, Italy

Reviewed by:
Elisabetta Lalumera,
Università degli Studi di Milano-Bicocca, Italy
Rita Temmerman,
Vrije Universiteit Amsterdam, Belgium
Pius Ten Hacken,
Universität Innsbruck, Austria

****Correspondence:***
Pamela Faber
pfaber@ugr.es

Specialty section:
This article was submitted to Language Sciences, a section of the journal Frontiers in Psychology

Received: *31 August 2015*
Accepted: *01 February 2016*
Published: *23 February 2016*

Citation:
Faber P and León-Araúz P (2016) Specialized Knowledge Representation and the Parameterization of Context. Front. Psychol. 7:196. doi: 10.3389/fpsyg.2016.00196

INTRODUCTION

According to Akman and Surav (1997) and Akman (2000), the denotation of *context* has become murkier as its uses have spread out in many directions to the extent that it has become a sort of 'conceptual garbage can.' For this reason, efforts are currently being made to parameterize and generally make sense of *context* and all that it implies. However, though instrumental in numerous disciplines, context has no universally accepted definition, because it can point to many different things. In the same way as the definition of any word, the definition of context can also vary depending on the field of application, such as Linguistics, Cognitive Science, or Computer Science (Bazire and Brézillon, 2005).

Specialized knowledge is related to all of these three areas in the sense that (1) it is shared and disseminated through linguistic communicative acts (journal articles, conferences, etc.); (2) it is processed and acquired in the mind; and (3) it may be subjected to formalization. Therefore, the

parameterization of context for specialized knowledge representation should be approached from a multidisciplinary perspective.

Specialized knowledge can be represented in a variety of formats (i.e., ontologies, vocabularies, thesauri, controlled languages, databases, etc.) that may or may not support context, because knowledge resources are conceived for very different purposes (i.e., classification, reasoning, knowledge acquisition, standardization, harmonization, information retrieval, machine, or human translation, etc.). More specifically, terminological knowledge bases (TKBs) generally describe the concepts and terms of specialized knowledge domains for users with linguistic and/or cognitive needs. TKB users are most often human (e.g., translators, experts, technical writers), but computer applications can also benefit from terminological resources when it comes to automatically interpreting and/or producing specialized texts. Even though TKBs usually provide conceptual representations based on some sort of knowledge modeling mechanism, they rarely support context modeling. In other words, very few provide controlled partial information concerning conceptual entities by viewing them from different viewpoints or situations. This can be a problem because the meaning, designation, collocates, and location of a concept within a knowledge configuration or linguistic structure often vary, depending on context.

Contextual information must thus be included in a TKB that aspires to being a knowledge representation resource. In this regard, it is timely and necessary to parameterize context in specialized knowledge domains with a view to more effectively facilitating knowledge representation, understanding, and acquisition. Nevertheless, matters are further complicated by the fact that context itself is a complex, multidimensional concept. Reasons for its conceptual fuzziness include the following: (i) there are various types of contexts; (ii) many types of data can be extracted from context analysis; (iii) contexts can also be used for a wide range of different purposes.

Contextual information can be specified in terms of scope (local vs. global) or according to the type of information conveyed (syntactic, semantic, and pragmatic variables). As reflected in corpus analysis, when context is mentioned in a text, it is metaphorically conceived as a container or a bounded space, since an utterance can be "in context" or "out of context." Context also *frames* or *surrounds* the utterance or object of analysis. In this sense, context bears a resemblance to Fauconnier's (1985, 1997) mental spaces since the location of an utterance in this bounded space or container is what makes it meaningful. As a relational construct in texts, context helps to anchor linguistic designations to objective reality by providing background information, situating objects and processes, and explicitly relating them to each other as well as to the agents that manipulate them and act on them. It is thus a constraining factor that drives understanding. In other words, as stated by Leech (1981), the specification of context (whether linguistic or non-linguistic) has the effect of narrowing down the communicative possibilities of the message as it exists in abstraction from context.

The remainder of this paper is organized as follows. The Section "What is Context?" reviews the notion of context as found in the literature of different areas. In Section "Context and Terminology," context representation is described with regards to terminology and specialized knowledge. The Section "Context Parameters" proposes a taxonomy of context parameters from a local to a global scope further divided into syntactic, semantic, and pragmatic facets. These facets cover the specification of different types of contextually relevant terminological information, such as predicate-argument structure, collocations, semantic relations, term variants, grammatical, and lexical cohesion, communicative situations, subject fields, and cultures. The examples given are drawn from the domain of environmental science based on the experience acquired while building EcoLexicon (ecolexicon.ugr.es), an environmental multilingual TKB. The Section "Conclusion and Future Work" provides the conclusions derived from the parameterization of context for specialized knowledge representation.

WHAT IS CONTEXT?

Research communities envision context differently since they conceive it in relation to different entities. Thus, context may be the parts of discourse surrounding a word, sentence, or passage, also known as co-text (Textual Linguistics), the set of situational elements where the object being processed is included (Cognitive Psychology), or that which surrounds and gives meaning to something else (Computer Science).

In Linguistics, context has long been regarded as an essential factor in the interpretation of linguistic utterances. It plays an important role in different tasks, such as meaning construction, inference, variation, modulation, sense disambiguation, etc. Quite often co-textual elements are sufficient to resolve ambiguity, but sometimes other context types also come into play.

Apart from the co-text sense, context in Linguistics is also mentioned in relation to pragmatic and cognitive notions, such as speech acts (Austin, 1962; Searle, 1969), conventions (Gadamer, 1995), maxims (Grice, 1975), framing (Goffman, 1974), common ground (Clark, 1996), and mutual manifestness (Sperber and Wilson, 1986, 1990), which refers to what one is capable of inferring or perceiving even if one has not done so as yet. The sum of these shared assumptions constitutes the cognitive environment of a group of individuals, which provides the foundation for successful communication (Yus, 2006).

These notions are related to sociocultural factors accounting for broader contextual variables, such as communicative settings, cultures, or world knowledge. Evidently, context has also been extensively studied in discourse studies, where it has been defined as the totality of conditions under which discourse is being produced, circulated and interpreted (Blommaert, 2005, p. 251). In the same area, Van Dijk (2005, p. 237) gives an even wider view by dividing context in different dimensions, namely, the cognitive, social, political, cultural, and historical environments of discourse.

In Cognitive Science, since the emergence of situated cognition, background situations have also become an essential element in the analysis of context. This has had an impact

on cognitive linguistics, where meaning is thought to be based mostly on situational context and constructed on-line (Croft and Cruse, 2004; Evans and Green, 2006). Meaning thus does not exist without context. For example, the theory of situated cognition argues that knowledge is situated, and is partly a product of the activity, context and culture in which it is developed and used (Brown et al., 1989, p. 32). Clancey (1994) adds that the situated aspect of cognition is that the world is not given as objective forms. Rather, what we perceive as properties and events is constructed according to the context.

Elman (2009, p. 572) highlights the importance of context in language comprehension and asserts that the meaning of a word is rooted in our knowledge of both the material and social world. Therefore, the meaning of a word is never 'out of context' even when we are not aware of what this context is. He also highlights the importance of larger knowledge structures: "events play a major role in organizing our experience. Event knowledge is used to derive inference, to access memory, and affect the categories we construct. An event may be defined as a set of participants, activities, and outcomes that are bound together by causal relatedness." Consequently, all lexical units, apart from their micro-context in discourse, need to be understood within the context of a larger event.

According to Yeh and Barsalou (2006, p. 350), knowledge of a larger event or situation restricts the entities and events likely to occur in it. Conversely, knowledge of current entities and events constrains the event or situation likely to be unfolding. Context thus plays a crucial role in knowledge understanding and acquisition since it can trigger one meaning while inhibiting another.

Cognitive processing necessarily includes linking an utterance or object to the right context, something that the human brain does with relative ease. In this sense, according to Flowerdew (2014), speakers and writers are remarkably adept at knowing which features of context to rely on to make their utterances meaningful, and listeners and readers are equally adept at contextualizing what they read or hear in order to understand it. However, what is not so easy is to agree on how to characterize context types and describe how they interact with each other. In fact, context was for a long time omitted in linguistic accounts because it was considered to be too chaotic and idiosyncratic, to be systematically characterized (Ervin-Tripp, 1996, p. 35).

Despite the evident challenge, the benefits of formalizing context are well known in computing. Computer Science has been dealing with context as a formal object –although more limited in scope– for some time now since McCarthy (1987, 1993), who stated that there is simply no most general context where all the stated axioms always hold and are meaningful.

From a computational perspective, contexts are useful for putting together a set of related axioms. In this way, contexts are used as a means for referring to a group of related assertions about which something can be said (Guha, 1991). However, the notion of context in computer science has two sides (Brézillon, 2005). Firstly, there is the cognitive science view, where context is used to model interactions and situations in a world of infinite breadth and human dimension, which is the key for extracting a model. Secondly, there is the engineering view, where context is useful in representing and reasoning about a restricted state space within which a problem can be solved. Since context, knowledge and reasoning are closely intertwined (Brézillon, 2005), the main aims of artificial intelligence with regards to the formalization of context seem obvious: (i) performing automatic inferences and reasoning (Guha, 1991; Lenat, 1995); (ii) identifying relational constraints for human–computer interaction and context-aware applications (Dey, 2001); (iii) improving automatic information retrieval, resolving ambiguities in natural language processing (NLP), etc. Also relevant to the parameterization of context is the concept of explanatory coherence (Thagard, 1989), which formalizes and computes coherence as a constraint satisfaction problem (Thagard and Verbeurgt, 1998).

Although some of these applications go beyond the scope of this proposal, there are others that could benefit from the systematization of context features in specialized knowledge resources, especially those related to NLP and domain ontologies. More specifically, with regards to knowledge representation and reasoning, context is needed to derive new knowledge from what is already known.

However, context is more than a set of previously specified discrete variables that have an impact on the knowledge of a language and a person's ability to use it. Context and language are considered to be in a mutually reflexive relationship, such that language shapes context as much as context shapes language (House, 2006).

CONTEXT AND TERMINOLOGY

As is well known, Terminology is the study of how specialized knowledge concepts are structured, described, and designated in one or various languages within a specialized domain. One of the practical tasks in Terminology is the design and creation of terminological resources so that users, whether human or artificial, can effectively access concepts and associated information in order to understand, acquire, or produce specialized knowledge.

Although the tendency in the General Theory of Terminology (Wüster, 1979) was initially to disregard context and contextual variables as well as the terminological variation that they produce, it soon became apparent that specialized terms are lexical items that are used in communicative contexts (Sager, 1990; Cabré, 1999), and that these contexts can affect their potential meaning. In fact, specialized knowledge units or terms acquire their meaning in context, more specifically, within a frame including a semantic and pragmatic background (Reimerink et al., 2010).

Nevertheless, contextual information is rarely found in specialized knowledge resources. As pointed out by Bowker (2011), most term banks present terms out of context, or in only a single context. A possible reason is the widespread belief that terms in the same field never have more than one meaning and thus have a one-to-one relation with the object or process designated. However, terms and concepts are dynamic and context-sensitive. For instance, concepts may be recategorized so as to constrain their relational behavior, and terms may show several types of variants with different cognitive, semantic, and

usage consequences (León-Araúz and Faber, 2014), (see examples in Sections "Local Pragmatic Contexts" and "Global Pragmatic Contexts").

User understanding of an entity or group of entities depends on having access to the necessary information to activate the right frame or knowledge structure in which the word or term should be processed. In turn, the effective production of a specialized utterance also depends on the user having access to the combinatorial potential of the terms involved. When a terminological resource includes multilingual correspondences, contextual information becomes even more crucial because of the lack of isomorphism between languages and cultures.

Generally speaking, even when contextual information is included in the concept or term entries of knowledge resources, it is not inserted in a systematic way since there is no consensus of opinion on the exact nature of context. The most common form of context found in terminological resources is a textual excerpt where terms are shown in real use, whether such contexts are in the form of KWIC (Key Word in Context) concordances or longer full-sentence segments. These can be useful to enhance both linguistic and cognitive user needs since they can provide valuable information regarding the collocational behavior of the terms and/or the relational behavior of the concepts activated. However, this is only a small fraction of what context representation should be.

Dubuc and Lauriston (1997, pp. 81–83) were among the first to underline the importance of context in Terminology: "Contexts are important to terminology with respect to the relationship of a term with its field of application. The context embodies the discourse bearing the term [. . .]. It is the presence of conceptual features relevant to the term that determines the extent of the context." Despite the fact that their interest in context was restricted to evidence of the term being used in the specialized field and the conceptual content associated with the term, this was still a relatively new assertion for the time. They classified contextual excerpts as associative, explicative, or defining, depending on the quantity and quality of concept descriptors obtained. This seminal study focused on context as a way of enhancing the reader's mental image of a concept.

Pearson (1998) goes somewhat further and explains why context is a great deal more than a text excerpt included in a term entry for purposes of knowledge acquisition. She affirms that the only way of determining what a term is and whether language is behaving 'terminologically' is by examining context. A context thus reflects a certain communicative setting, which is the most important factor that shows whether a given lexical unit is being used as a term or as a general language word. Finally, she also highlights the usefulness of metalanguage patterns retrieved from corpora in the formulation of terminological definitions (ibid: 191–203). This was subsequently complemented by Meyer (2001), who introduced the notion of *knowledge-rich context*.

Not surprisingly, in the last 15 years, context has become an important focus in Terminology research and its uses have multiplied accordingly. In its co-text sense, it is currently a primary data source for elaborating and constraining the scope of meaning definitions. It has thus become a rich source of complementary conceptual information, linguistic usage, and knowledge representation, *inter alia*. Nevertheless, as observed in Section "What is Context?," context encompasses much more. Other than static text-based usage examples, context representation in Terminology should also cover background situations, cultures, communicative settings, etc.

The vital role of specifying context and of embedding specialized concepts in situations has been highlighted as a way of enriching conceptual representations in TKBs. According to Meyer et al. (1992), TKBs should reflect conceptual structures similarly to how concepts are related in the human mind. Similarly, Faber (2011) states that the organization of semantic information in the brain should underlie any theoretical assumption concerning the retrieval and acquisition of specialized knowledge concepts as well as the design of specialized knowledge resources.

For example, in an fMRI study of expert-novice differences in the identification of geological field instruments, Faber et al. (2014b) found that in contrast to novices, experts activated the bilateral precuneus, posterior cingulate, and insula, three regions previously implicated in mental imagery, episodic memory, and context representation. In addition, the importance of visual scene generation was reinforced by brain activation in the parahippocampal gyrus, which encodes meaningful contextual associations.

In Frame-Based Terminology (FBT; Faber et al., 2005, 2006, 2007; Faber, 2011, 2012, 2015), specialized knowledge units are only understood with reference to their underlying conceptual frame, whose elements are selected according to context. Context determines the activation of previously stored knowledge and the formation of new categories (Croft and Cruse, 2004 p. 75). In this sense, Barsalou (1983, 1991) found that conceptual categories can be created in an *ad hoc* goal-derived way, which indicates that context determines the conceptual organization underlying a concrete situation. Since categorization itself is a dynamic context-dependent process, the representation and acquisition of specialized knowledge should certainly focus on contextual variation (León-Araúz et al., 2013).

For this reason, one of the keys to the enhancement of specialized knowledge resources lies in parameterizing contextual information. This entails distinguishing different types of context, their scope and facets as well as how they interact with each other. This is not a simple objective to achieve despite the fact that specialized discourse does not have as many contextual variables as those in general language (e.g., figurative meaning, irony, etc.).

A solid theory of context and context types would be a timely contribution to lexical semantic research which would have repercussions in a wide range of fields. A principled set of context modeling parameters would facilitate knowledge acquisition and understanding. Such resources would ideally allow non-experts to understand a given domain by focusing on and capturing essential knowledge. However, they would also benefit diverse applications in NLP and in the Multilingual Semantic Web (MSW; León-Araúz and Faber, 2014). The MSW is envisioned as an information space where language-independent knowledge would be accessible across different natural languages. This entails the improvement of many NLP techniques related to both comprehension and production,

such as word sense disambiguation, cross-lingual mappings, or question answering –always depending on general language resources such as WordNet. Thus, for the web to be truly semantic and multilingual, different NLP tools and techniques need to rely on high-quality multilingual resources –whether general or specialized– that account for the representation of context, a major barrier to successful communication.

CONTEXT PARAMETERS

Many authors have proposed the characterization of context types, based on a wide range of different criteria. In Cognitive Linguistics, Evans and Green (2006, p. 21) underline the importance of different types of context in the modulation of any given instance of a lexical item as it occurs in a particular usage event. Broad context types mentioned are the following: (1) encyclopedic context (information accessed within a network of knowledge); (2) sentential context (utterance meaning); (3) prosodic context (intonation pattern); (4) situational context (physical location where the text is emitted); and (5) interpersonal context (relationship holding between text sender and receiver). Most other approaches give a more binary vision of context. For instance, Harris (1988) proposes *world knowledge* vs. *language knowledge*, whereas Halliday (1989) makes the distinction, *context of situation* vs. *context of culture*. This duality can also be found in the distinction between context and co-text.

In reference to specialized knowledge units, the primary division of context is based on scope, since contexts can be either local or global. Context may be a few words on either side of a term (He et al., 2010), the sentence or paragraph in which it is appears (Soricut and Marcu, 2003), a set of documents containing it (Cilibrasi and Vitanyi, 2007), a communicative act, or even a whole culture. According to Akman and Bazzanella (2003, p. 325), an adequate multi-modal coding of context on both the global and local levels would be useful in delimiting inferences, disambiguating deictic expressions, and solving the problem of indeterminacy.

Thus, the distinction of local vs. global can be found elsewhere in the literature though not with the same meaning. Bazzanella (1998), Akman and Bazzanella (2003), and Miecznikowski and Bazzanella (2007) refer to local context to denote a specific setting where the participants interact; and use global context for referring to the members of a community, their social norms, culture, beliefs, ideology, etc. In the same way, Mihalcea (2007) uses the same distinction to refer to a different context span within textual excerpts (a pair of words vs. lexical chains), whereas Dash (2008) proposes a continuum of four contexts from local to global: (i) local context (the immediate environment of a word); (ii) sentential context (syntactic-based); (iii) topical context (domain-based); and (iv) global context (extralinguistic reality).

In our view, local contexts are usually limited to the words within the term itself, to a small number of words in the immediate vicinity of a term, or to words connected by syntactic dependencies to the term. According to Agirre and Stevenson (2007, p. 225), the data that can be derived from local contexts are the following: part of speech, morphology, collocations, subcategorization, frequency of senses, syntagmatic and paradigmatic word association, selectional preferences, semantic roles, domain, topical word association, and pragmatics. Evidently, these categories are not watertight containers since there is a great deal of overlap between information types but they are all valuable data categories to be included in a TKB.

In contrast, global contexts can encompass the whole text or go beyond the text: to the communicative situation (i.e., formal vs. informal); to the conceptual networks reflected in it; to the culture in which the text is interpreted, etc. This means that global contexts refer to items that are often quite a distance from the term or even outside of the text altogether though within the specialized domain.

Both local and global contexts can be subdivided, based on whether they are mainly syntactic, semantic, or pragmatic. In our opinion, it is extremely difficult to trace a clear boundary line between syntax, semantics, and pragmatics because there is a significant degree of overlap. In fact, in Cognitive Linguistics, the distinction between semantics and pragmatics is even rejected. For example, at the local level, the use of term variants *drainage basin* or *catchment area* instead of *river basin* or *watershed* (all expressing the same concept) has pragmatic significance since it signals that the text sender has expert knowledge and is British or Australian instead of American. However, the choice of *drainage basin* also has a semantic dimension since *drainage* foregrounds water movement and accumulation which are the processes that occur in this area whereas *river basin* only foregrounds the location of the basin without any reference to water flow.

At the same time, the term also possesses a syntactic dimension. *Drainage* modifies *basin*, the head of the multi-word term. The implicit relation between modifier and head can be expressed by the preposition *for* (basin for drainage) since the basin is where drainage occurs. However, the structure of the term can be unpackaged to *basin where water drains in and then drains out*. It is thus the result of meaning compression given the fact that *drainage* encodes both the incoming and outgoing flow of water.

This interaction reflects the fuzzy boundaries between syntax, semantics, and pragmatics in general and specialized language. As Dash (2008, p. 29) points out, each context is interlinked with the other in an invisible thread of interdependency. This fuzzy three-level approach to context goes hand in hand with the micro-theories proposed by FBT, which are related to the information encoded in term entries, the relations between specialized knowledge units and the concepts that they designate (Faber, 2015, p. 15).

Local Contexts

Local contexts are generally regarded as spans of +5 items before and after the term occurrence. They are important in the design stage of a TKB for a wide variety of reasons, which include (but are not limited to): (i) term disambiguation; (ii) meaning definition formulation; (iii) specification of linguistic usage; (iv) conceptual modeling; and (v) term extraction. Thus, local contexts can be used either by resource creators in order to

develop terminological by-products (i.e., definitions, conceptual networks, usage examples, etc.); or by the users themselves (i.e., obtaining direct access to the corpus).

In Corpus Linguistics, a recurring local context is known as a *collocation*. However, *collocation* is a rather vague term that does not cover the same range of linguistic phenomena for all linguists (Mollin, 2009, p. 176). The definition of a collocation, its length, the neighborhood of possible collocates and their strengths of occurrence (*inter alia*) are all part of the analysis of specialized language texts and the terms that they contain. Needless to say, the representation of this type of information should be an important element in the design of data fields in terminological entries. However, there are different ways to approach collocational information: they can be seen as a combination of grammatical elements, as the codification of semantic relations, or as pointers to pragmatic information.

Therefore, such local contexts can be syntactically parameterized based on syntagmatic patterns and/or semantically mapped in terms of the interaction, foregrounding, or specification of the definitional features of the concepts activated in them. As shown in Section "Local Pragmatic Contexts," certain types of pragmatic information are also reflected in local contexts. This occurs, for instance, when term variants indicate changes in the knowledge area, specialization level, geographic region, cultural community, and/or historical period.

Local Syntactic Contexts

Local syntactic contexts are those that reflect the recurrent structural patterns in which the term participates. Terms have a combinatorial value and distinctive syntactic projections. However, a term's position in a subject or direct object slot or as the head of a prepositional phrase is often not very informative since the fact that a term has a certain grammatical function in a sentence is not always relevant to its meaning. Nor is the analysis of a multi-word terminological unit as a mere combination of grammatical categories much more helpful unless this pattern is linked in some way to its underlying semantics. It is more productive to take a semantic view of syntax and to analyze syntactic contexts as the linguistic codifications of predicate-argument structure.

In this regard, each predicate can be said to have an argument structure or valence, specifying the number of arguments that it can take. The concept of valence was first proposed by Tesnière (1959) and now plays a crucial role in the majority of today's linguistic theories. Generally speaking, valence is regarded as the ability of certain lexical units (e.g., verbs) to open slots which are filled by other lexical units. Valence can be envisaged syntactically, semantically or as a combination of the two. Again, this is proof of the fuzzy interaction of context types and parameters.

A predicate's valence depends on its meaning since its arguments are essentially the participants which are minimally required for the activity or state described. Such representations should thus include the decomposition of the predicate and the specification of the semantic characteristics of the arguments (Faber and Mairal Usón, 1999).

Despite the fact that verbs have never been a primary focus in Terminology, approaches to syntax in Terminology can benefit greatly from linguistically sensitive theories of lexical structure that focus on verbs and on how their meaning relates to syntactic forms within a sentence. One reason for this is that verbs play an important role in specialized discourse because their position in a lexical domain and degree of semantic specificity is in direct relation to the number and type of arguments that they can combine with (Faber and Mairal Usón, 1999). In specialized texts, these arguments are terms or specialized knowledge units, whose semantic characteristics constrain the polysemy of the verb and even model its meaning. In this sense, one can say that the meaning of general language verbs can be significantly modified or even transformed by their context of activation. When general language verbs appear in domain-specific texts, they become specialized because their arguments constrain their meaning (L'Homme, 2003). At the same time, the presence of a particular verb also constrains the type of argument slots that specialized terms may fill.

For example, *dissipate* is a polysemic general language verb, which is often found in scientific discourse. When it is used transitively in the sense of one entity dissipating another entity, it has two arguments. The first argument has the semantic role of *agent* and the second has the role of *theme*. In this regard, the argument structure of *dissipate* is fairly straightforward since X (agent) causes Y (a theme undergoing the action) to be dissipated:

(1) Dissipate $(x)_{agent}$ $(y)_{theme}$

According to the *Merriam–Webster Dictionary*, this transitive use of *dissipate* has one of the following four senses: (i) to break up and drive off (as a crowd); (ii) to cause to spread thin or scatter and gradually vanish; (iii) to lose (as heat or electricity) irrecoverably; (iv) to spend or use up wastefully or foolishly. Contextual data extracted from the enTenTen12 general English corpus in *Sketch Engine* show that *dissipate* is often used unaccusatively. In other words, the first argument is not made explicit. The most frequent meanings of *dissipate* in general language are ii (2) and iv (3):

(2) *To cause to spread thin or scatter and gradually vanish*

- (i) Temperature (e.g., *warmth, heat*). [When you exercise on land, sweat evaporates, and cools your skin to **dissipate** *heat.*]
- (ii) Meteorological phenomena (e.g., *storm, fog, mist*). [By afternoon, however, the air traffic from the city had become normal again when the *fog* **dissipated** almost completely.]
- (iii) Visual/olfactory perception (e.g., *mirage, smell*). [I hope most of the *smell* **dissipates** by the time that I ride my bike this afternoon.]
- (iv) Emotions/feelings (e.g., *fear, anxiety*). [With the knowledge, the *anger* **dissipated** as quickly as it had come.]

(3) *To spend or use up wastefully or foolishly.*

(i) Valuable possessions (e.g., *wealth, resources*). [The entrepreneurs have become wealthy while showing how extravagance and luxuries **dissipate** *wealth.*]

As can be seen in the case of sense (2), the dissipated entity is most frequently related to temperature, weather, sensory perception, and emotions, whereas in sense (3), it is generally wealth or financial resources.

However, in specialized contexts, the meaning of *dissipate* does not really correspond to any of these possibilities. The reason for this is the semantic content of domain-specific arguments, which interact with the base meaning of *dissipate* and model it to create a new sense that is apt for scientific contexts. In the EcoLexicon corpus (subdomain Coastal Engineering), the statistically significant collocates of *dissipate* in the theme slot are the following: (i) *energy* (e.g., *flux, gradient, power*); (ii) *cyclone* or a storm-related term (*tide, wave, wind,* etc.), which can also be regarded as a type of energy. More specifically, energy appears as the generalization of heat whereas cyclone is a specification of energy. **Table 1** shows examples of specialized contexts for *dissipate.*

As can be observed, the arguments of *dissipate* are all NPs that belong to the same semantic categories and combine in similar patterns. In (4) in **Table 1**, the first argument is a process involving some type of friction and the second argument is energy. In (5) in **Table 1**, the storm entity appears unaccusatively without explicitly referring to the reasons for energy loss, which reflects the fact that the target audience is already aware that reasons for storm dissipation include colder sea surface temperatures, shearing winds, sinking air, moving over land, depending on their type, location, and intensity.

In both cases, the interaction of the semantic characteristics of both the dissipated entity (*energy*) and dissipating process (*friction, breaking, falling, uprushing*) clearly point to a new (specialized) sense of *dissipate*, which responds to the Laws of Thermodynamics:

(6) To cause (energy) to be lost through its conversion to heat.

The definition of *dissipate* in (4) fits the domain of Coastal Engineering. The energy produced by wave movement is *dissipated* (lost), typically from friction or turbulence when the waves are near the shore and come into contact with the sea bottom. Of course, the energy is not actually lost but rather is transformed into heat, which raises the temperature of the system. The conversion to heat, though explicit in the definition, is not lexicalized in contexts since it is part of the shared knowledge in the domain. This is one example of how verbs within domain-specific contexts become transformed when they are used in specialized texts since the terms that fill the slots in their argument structure contextualize, modify, and/or restrict their meaning. Moreover, since arguments are specialized terms and verbs are relational constructs, the analysis of argument structure can lead to the construction of semantic networks or frames, which again reflects the fuzzy boundaries between syntax and semantics. For instance, all arguments in the second specialized sense of *dissipate* are cohyponyms (*tropical cyclone*, *hurricane*, *tornado*).

Another way of viewing a local syntactic context is as a *colligation*, initially defined as the co-occurrence of grammatical categories (Firth, 1968, p. 181) or the grammatical company that a word keeps. In multi-word expressions (MWEs), the relations between words and their grammatical categories cover a wide spectrum. In most cases, the words are linked by both grammatical and lexical relations. In fact, it is difficult if not impossible to determine which relation is stronger in each case.

According to Hoey (2005, p. 43), the basic idea of colligation is that in the same way that a lexical item can be primed to occur with another lexical item, it can also be primed to occur in or with a particular grammatical function. Colligation is concerned with the typical grammatical patterning of words (or word classes). As such, collocation and colligation are not totally separate concepts, but together create a network of meaning. Distinguishing collocations (co-occurrences of words) from colligations (co-occurrence of word forms with grammatical phenomena; Gries and Divjak, 2009) is not always a simple task. There is no clear boundary between various types of word combinations inasmuch as they can be simultaneously a collocation and a colligation.

This highlights the interdependence of syntax and lexis. For example, whereas V + *out* + NP is a colligation, *spew* (V) + *out* + *air pollution* (NP) is a collocation which exemplifies the colligation. However, the meaning of *colligation* has since expanded to include the specification of semantic preference or semantic prosody. Accordingly, it can now refer to the co-occurrence of lexis and grammatical categories.

Semantic preference is the "relation between a lemma or word-form and a set of semantically related words" (Stubbs, 2002, p. 65). Semantic prosody (Louw, 1993) captures the fact that some elements attract lexical items designating negative things, features, actions, etc., whereas others show a characteristic co-occurrence with positive elements. Together these notions expand into the notion of semantic colligation: the mutual attraction holding between a grammatical construction and a semantic category (Gabrielatos, 2007).

For example, when something is projected (*spewed, pumped*, etc.) from a container, the ejected entity (*air pollution,* CO_2) is frequently undesirable. Moreover, this semantic preference is confirmed by corpus data showing that when *pollution* is the theme argument, it tends to consistently combine with verbs belonging to two semantic domains:

(i) Change: to cause to become less (*reduce, decrease, abate, minimize, mitigate, lessen, cut*).
(ii) Movement: to move out of/from a certain place (*spew, emanate* with the focus either on the endpoint of the trajectory (*pollution spewed into the air/environment*) or its starting point (*pollution spewed from the Chinese mine/coal-fired power plant/plane*).

When *pollution* or one of its components is the agent argument, it also tends to combine with verbs of change, but primarily with those predicates belonging to the subdomain *to cause something to become worse* (*contaminate/foul/degrade*, etc.).

(7) Still more fluorides from such pollution ***contaminate*** the animals and plants we use as food.

TABLE 1 | Contexts of *dissipate* in specialized texts.

Concordances	Pred-arg structure
(4) Energy (*energy*)	
The wave *energy* has been **dissipated** by wave *breaking* and *bottom friction*.	Dissipate (wave breaking and bottom friction)$_{agent}$ (energy)$_{theme}$
Part of the *energy* is **dissipated** by *breaking processes*.	Dissipate (breaking processes)$_{agent}$ (energy)$_{theme}$
X is the fraction of *energy* **dissipated** by the *falling* sand grains.	Dissipate (falling sand grains)$_{agent}$ (energy)$_{theme}$
Part of the wave *energy* is **dissipated** by the *uprushing* water body.	Dissipate (uprushing water body)$_{agent}$ (energy)$_{theme}$
(5) Meteorological phenomenon (e.g., *cyclone, hurricane, tornado*)	
Only if the *tropical cyclone* **dissipates** with just a tropical disturbance remaining with the OAR give the system a new name.	Dissipate (ø)$_{agent}$ (tropical cyclone) $_{theme}$
Hurricanes **dissipate** when their energy supply is substantially reduced.	Dissipate (ø)$_{agent}$ (hurricane) $_{theme}$
Even though the *tornado* is **dissipating**, the tornado is still capable of causing damage.	Dissipate (ø)$_{agent}$ (tornado) $_{theme}$

(8) The measure should help rescue the Chesapeake Bay from the nutrient pollution ***fouling*** its water.
(9) Changes in the flow of the river and pollution have very severely ***degraded*** the ecosystem.

Optionally, it is also activated with verbs of causative existence in the subdomain of *to cause something bad to happen.* More specifically, *threaten* means *to cause something or someone to be vulnerable or at risk.*

(10) The oceans are now so ***threatened*** by pollution and exploitation that many shorelines will soon be totally denuded of marine life.

The predisposition to appear in certain syntactic structures and combine with predicates from specific semantic subdomains is directly related to the semantic load of *pollution,* a dot object according to the (Generative Lexicon) (Pustejovksy, 1995), which can be regarded either as a process or the result of a process.

Both colligations and predicate argument structures reflect the fuzzy boundary between syntax and semantics. The fact that words that occur together tend to be semantically similar explains that local syntactic contexts could also be used for semantic clustering. Popularized by Firth (1957) in his famous line "a word is characterized by the company it keeps," this approach has been implemented in the distributional hypothesis (Harris, 1985) and the strong contextual hypothesis (Miller and Charles, 1991). Furthermore, within the scope of a sentence, word sense disambiguation is usually determined by a combination of two factors: (1) the syntactic frame into which the word is embedded, and (2) the semantics of the words with which it forms syntactic dependencies (Rumshisky, 2008, p. 217).

Local Semantic Contexts

Local semantic contexts can either refer to semantic relations between the constituents of the specialized knowledge unit (term-internal semantic context) or to semantic relations between different specialized knowledge units in the text (term-external semantic context). In the first case, the scope of the context is the multi-word term itself, whose interpretation is based on the meanings as well as the dependency relations between the head and the modifiers from which a semantic relation can be inferred. In the second case, the context is the linguistic codification of a triplet: two specialized knowledge units linked by a phrase that explicitly marks a semantic relation.

Term-internal semantic contexts are exemplified by MWEs and constitute a large portion of the lexicon of any natural language. It is estimated that the number of MWEs in the lexicon of a native speaker is the same as the number of single words (Jackendoff, 1997), and these ratios are probably even higher in the case of domain-specific language, in which the specialized vocabulary and terminology are composed mostly of multiword expressions. According to Erman and Warren (2000, p. 29), the fact that half of spoken and written language comes in pre-constructed multiword combinations makes it impossible to consider them as marginal phenomena. In fact, these specialized MWEs are rapidly increasing because of the continuous addition of new terms that designate new concepts. This makes it virtually impossible to store all of them in a dictionary.

Since the meaning of multi-word terms is often a specialization of the meaning of its head, in many cases, term structure can be used as a way to automatically extract information for the specification of conceptual hierarchies, one of the main components of TKBs. In morphologically poor languages, such as English, they can take the form of sequences or stacks of nouns of varying length: (i) two constituents (*capillary wave*); (ii) three constituents (*long-wavelength surface wave*); (iii) four constituents (*ocean surface gravity wave*); and (iv) five constituents (*surface gravity wave elevation spectra*). It is for the addressee to unpack their meaning and determine the implicit relationship between the constituents. One way of doing this is to understand such compounds in terms of left or right branching dependency relations:

(11) [capillary (wave)]
(12) [long-wavelength (surface wave)]
(13) [ocean (surface gravity wave)]
(14) [((surface gravity wave) elevation) spectra]

Evidently, the more numerous the terms in the stack, the more difficult it is for a computer to automatically establish dependencies. The most complicated cases are (13) and (14). For example, in (13), it is necessary to know that the term must be interpreted as a right-branching compound, since the term does not refer to an ocean surface. Instead, *surface gravity*

wave designates an important type of wave (which is a synonym of *surface wave),* which propagates in the ocean. In contrast, (14) is a left-branching compound, which is gradually generated from *surface gravity wave* by adding the subsequent heads of *elevation* and then *spectra.* The analysis of these dependencies is often based on term entrenchment and extremely difficult for a computer to perform automatically unless it has been previously trained to do so.

Another way of understanding term creation is to think of the general concept as an entity with an underspecified meaning. Its nature is what predetermines the potential specification of its meaning that will lead to the generation of hyponyms.

For example, *wave* designates an oscillation propagating through a medium. As such, it has a series of basic defining attributes, such as height, wavelength, steepness, period, speed, and frequency. This can be regarded as a kind of frame or type of implicit context that opens up slots. The specification of one or more of these attributes that modify the head (*wave)* creates different wave types (*long-period wave, high-frequency wave*, etc.).

However, these defining attributes are not the only source of new terms. When a wave is regarded as a process, this creates a more abstract context or frame that opens the door to other possible subtypes. A wave can thus be regarded as caused by an agent (*wind wave*), affected by a force (*gravity wave*), moving in a certain way, (*plunging breaking waves*), taking place at a certain location (*surface wave*), and occurring in a certain medium (*water wave).* However, in some cases, both scenarios or contexts can combine to produce terms such as *long-wavelength surface gravity wave*. This term has the following meaning relations:

(15) Long-wavelength surface gravity wave
Head: wave
Located_at: surface
Affected_by: gravity
Length_of : long-wavelength

A long-wavelength surface gravity wave can thus be regarded as an oscillation (*wave*) with a long wavelength (*long-wavelength*) on the air–sea interface (*surface*) affected by a restoring force (*gravity*).

Evidently accessing the meaning of such a combination is not a trivial process since it activates a whole specialized frame that requires previous knowledge. In this sense, Maguire et al. (2010, pp. 49–50) cite the concept specialization model (Murphy, 1988) and dual process theory (Wisniewski, 1997), which propose a two-stage interpretation process. The first stage involves a slot-filling mechanism where the modifier is inserted into a slot in the head-noun schema to form an interpretation. The type of noun modifier is directly related to the basic meaning of the head.

For instance, in $N + N$ compounds in which *energy* is the headword, the slot activated is usually agent (e.g., *wave energy, wind energy, heat energy*, etc.), which highlights the source or natural force producing the energy. In contrast, in sediment compounds, in most cases, the headword opens a <location> slot (e.g., *intertidal zone sediment, streambed sediment, aquifer sediment*, etc.) since sediment is solid fragmented material that is transported and deposited by water or wind at a certain location. Alternatively, there is also a set of sediment terms in which a <made-of> slot is opened up (*lithogenous sediment, biogenous sediment, hydrogenous sediment, cosmogenous sediment*). These A + N compounds foreground the "fragmented material" part of the definition.

Consequently, this indicates that the membership of the head noun in a small number of broad semantic categories reveals consistent patterns in modifier and head use, and that the semantic categories are not randomly paired.

Term-external semantic contexts take the form of KWIC concordances or knowledge-rich contexts that provide information about a concept's attributes or the relations that it forms with other concepts. They contain Knowledge Patterns (KPs; Barrière, 2004), which are lexico-syntactic patterns that indicate a semantic relationship, and at least two specialized knowledge units. Studies in this tradition include Pearson (1998), Meyer (2001), Barrière (2004), Aussenac-Gilles and Jacques (2006), and Sierra et al. (2008), *inter alia*.

Mitkov (1998, 2002) and Meyer (2001) distinguish between knowledge-rich contexts and knowledge-poor contexts. Knowledge-poor contexts do not include any item of domain knowledge related to the search word. In contrast, knowledge-rich contexts contain at least one item of domain knowledge that is useful for the conceptual analysis of the search word. Such contexts should indicate at least one conceptual characteristic, whether it is an attribute or relation (Meyer, 2001, p. 281).

For example, the concordances of *erosion* in **Figure 1** show how different KPs convey different relations with other specialized concepts. The main relations reflected in erosion concordances are *caused_by*, *affects*, *has_location,* and *has_result*, which highlight the procedural nature of the concept and the important role played by non-hierarchical relations in knowledge representations.

In **Figure 1**, erosion is related to various types of agent, such as storm surge (1, 7), wave action (2, 13), rain (3), wind (4), jetty (5), construction projects (6), mangrove removal (8), surface runoff (9), flood (10), human-induced factors (11), storm (12) and meandering channels (14). They can be retrieved thanks to all KPs expressing the relation *caused_by*, such as *resultant* (1), *agent for* (2, 3), *due to* (6, 7), *responsible for* (11), and *lead to* (13). This relation can also be conveyed through compound adjective phrases, such as *flood-induced* (10) or *storm-caused* (12) and any expression containing *cause* as a verb or noun: *one of the causes of* (9), *cause* (4, 5, 8), and *caused by* (14).

Erosion is also linked to the patients it *affects*, such as water quality (15), sediments (16), coastlines (16), beaches (17), buildings (18), deltas (19), and cliffs (20). However, the affected entities, or patients, are often equivalent to locations (e.g., if erosion *affects* beaches it actually *takes place at* the beach). The difference lies in the KP linking the propositions. The *affects* relation is often reflected by the preposition *of* (10) or by verbs such as *threatens* (18), *damaged by* (17) or *provides* (19). In contrast, the *has_location* relation is conveyed through directional prepositions (*around*, 21; *along*, 22; *downdrift*, 23) or spatial expressions, such as *takes place* (24). In this way, erosion is linked to the following locations: littoral barriers (21), coasts (22), and structures (23). *Result* is an essential dimension in the description of any process since it is not only initiated by an agent

affecting a patient in a particular location, but also has certain effects, namely, the creation of a new entity (sediments, 25; primary coasts, 26; beach material, 27; shorelines, 28; marshes, 29; bays, 31) or the beginning of another process (seawater intrusion, 31; profile steepening, 32).

As can be seen, all these related concepts are quite heterogeneous. They belong to different paradigms in terms of category membership and/or hierarchical range. For instance, some of the agents of erosion are natural (wind, wave action) or artificial (jetty, mangrove removal) and others are general concepts (storm) or very specific ones (meandering channel). This explains why knowledge extraction must still be performed manually or semi-automatically and how local semantic contexts can be conceptually valuable. Nevertheless, it also illustrates one of the major problems in knowledge representation: *multidimensionality*. *Multidimensionality* has been defined by

```
Caused_by

1       , Alabama. Significant storm surge and resultant beach erosion were  associated with Ivan's landfall. However,
2       nd climate  On the Castellón coast, the main agent for erosion is wave action, and this is therefore  responsi
3       f a stream. The first   factor, rain, is the agent for erosion, but the degree of erosion is governed by   oth
4       rts (BW) and semiarid steppe (BS). Wind can also cause erosion and deposition in environments where sediments
5       etty. Reflection of waves from a jetty may  also cause erosion of adjacent shorelines. However, erosion furthe
6       oastal zone management. However, in some cases coastal erosion\par  can be due to construction projects that a
7       tude of about 0,3 M m3 per year.  Acute erosion  Acute erosion due to storm surges  (waves and water levels at
8       er. Mangrove removal is also reported to cause coastal erosion and change sedimentation patterns and shoreline
9       [edit] Erosion  Surface runoff is one of the causes of erosion of the earth's surface. Reduced   crop producti
10      pes. Local disturbances, for instance by flood-induced erosion, redistribution of sediment or accumulation of
11      ors and human-induced factors  responsible for coastal erosion and highlight the time and space patterns withi
12      ocess is typical of a cyclical process of storm-caused erosion in winter, followed by progradation\par  owing
13      can cause excessive wave action that can lead to beach erosion. Trash dumped from boats can be washed up onto
14       that have reached base level develop broad valleys by erosion caused by meandering channels. The stream chann

Affects

15      ing these sensitive creatures.  In some cases, coastal erosion can have adverse effects on water quality and h
16      ine Depositional Coasts  The erosion of coastlines and erosion of sediments being carried to the   shoreline b
17       use of dredged material to restore beaches damaged by erosion. EPA works with the U.S. Coast Guard to regulat
18      reasonable points, though when push comes to shove and erosion threatens buildings, traditional beach maintena
19      ks and arches found on irregular rocky coastlines; and erosion provides the material which forms deltas and ba
20       near the base of the cliff. Constant undercutting and erosion causes the cliffs to retreat landward.

Has_location

21      ed by the position of sand accumulation\par  and beach erosion around littoral barriers. A coastal structure i
22      hes. Kuenen (1950) estimates\par  that beach and cliff erosion along all coasts of the world totals about 0.12g
23      ce and divergence of wave energy over an offshore bar, erosion downdrift of a structure such as a groin, sudde
24      proportional to the longshore transport rate, and\par  erosion takes place downdrift at about the same rate. T

Has_result

25       Excessive loads of silt and other sediments caused by erosion can suffocate bottom-dwelling plants and animal
26      islands or coral reefs.  Primary coasts are created by erosion (the wearing away of soil or rock), deposition
27      \par  transported. Beach material is also derived from erosion of the coastal formations caused by\par  waves
28      ed to the   passage of the ice. Shorelines produced by erosion of glacial till deposits differ   markedly from
29      beaches   and marshes, are being formed as a result of erosion   and transportation of unconsolidated material
30      ion of the seashore and a rise in SLR. The results  of erosion could lead to further seawater intrusion that c
31      fs are developed  in landslide debris. In this cliffs, erosion of softer material has created bays. The  expec
32      s of steep systems, a sea-level rise may cause coastal erosion resulting in profile steepening, and therefore
```

FIGURE 1 | Non-hierarchical relations associated with EROSION.

```
Is_a

33      vided by the  area (A) of the drainage basin       (1) Erosion is the natural process of removal of soil by wa
34      in the Netherlands, geomorphological processes such as erosion, transport and sedimentation of sandy materials
35      BURY AND DUXBURY, 1996).     Coastal processes such as erosion and accretion are site-specific, season-specifi
36      these  catchments include:  Stormwater impacts such as erosion, channelisation, sediment deposition and sedime

Type_of

37      eroded by shallow overland flow (sheet, rill and gully erosion) and delivered to the drainage network. Channel
38      m the great   local relief, the result of differential erosion by glacier ice. Figure 9-20   includes two sche
39      ing flood events, the dikes are subject to the lateral erosion of the river   trying to reoccupy its former co4
40      d enlarges these small channels and generates headward erosion directed towards the aggrading active channel (
41      out five percent of the material on most beaches. Wave erosion of rocky coasts is usually slow, even where the
42       of the Earth´s land surface is   dominated by fluvial erosion, lakes that do occur are threatened with either
43      ind   climate, topography and surface roughness.  Wind erosion risk applies only when soils are dry and not co
44      oportional to the steepness of the land surface. Water erosion is in proportion to the shear stress exerted by
45      ley to become both wider and deeper over time. Glacial erosion also results in a change in the valley's cross-
46      dominate in periglacial environments: nivation; eolian erosion and deposition; and fluvial erosion and deposit
47       erosion processes. 215 CHAPTER 13 EQUATIONS: SEDIMENT Erosion caused by rainfall and runoff is computed with
48      gineers to simulate  cross-shore beach, berm, and dune erosion produced by storm waves and water levels. The l
49      uctures constructed to date have resulted in shoreline erosion   in their lee. Furthermore, the key environmen
```

FIGURE 2 | Hierarchical relations associated with EROSION.

many authors (Bowker, 1997; Kageura, 1997; Wright, 1997; Rogers, 2004) as the phenomenon in which certain concepts can be classified according to different points of view or conceptual facets. Evidently, multidimensionality has important consequences regarding how domains are categorized and modeled (León-Araúz et al., 2013). This is better exemplified in the concordances shown in **Figure 2** since multidimensionality is most often codified in the *is_a* relation.

In the scientific discourse community, concepts are not always described in the same way because they depend on perspective and subject-fields. For instance, *erosion* is described as a natural process of removal (33), a geomorphological process (34), a coastal process (35), or a stormwater impact (36). The first two cases can be regarded as conventional ontological hyperonyms. The choice of one or the other depends on the upper-level structure of the representational system, its level of abstraction and the support for context. However, the other two cases (i.e., coastal process and stormwater impact) cause the concept to be framed in more concrete subject-fields and referential settings.

The multidimensional nature of erosion is also clearly shown in subtypes, which are codified in term-internal semantic contexts. Erosion can thus be classified according to the dimensions of *result* (sheet, rill, gully, 37; differential erosion, 38), *direction* (lateral, 39; headward erosion, 49), *agent* (wave, 41; fluvial, 42; wind, 43, 46; water, 44; glacial erosion; 45), and *patient* (sediment, 47; dune, 48; shoreline erosion, 49).

These dimensions are contexts that need to be specified in the TKB in order to delimit information retrieval and make it more relevant. They can be represented as part of a definitional template (all cohyponyms being defined according to the same dimensions). Alternatively, they can be codified as a specification of the subsumption relation (fluvial erosion *is_a (agent)* erosion), or simply as a concordance or knowledge-rich context.

In order to retrieve new related term pairs, KPs can be collected and systematized in the form of local grammars. For instance, **Figure 3** shows part of the formalization of the causal relation, which is based on causative verbs in any of their inflected forms (*cause*, *produce*, *generate*, *trigger*), morphological particles (*-driven*, *-induced*) and other literal causative expressions (*responsible for*), as exemplified in the concordances of erosion (**Figure 1**). When this grammar is applied to the corpus, it identifies structures such as "tsunamis, usually caused by large earthquakes" or "rain produced severe flooding," from which we can derive the conceptual propositions, or triplets, TSUNAMI *causes* EARTHQUAKE and RAIN *causes* FLOODING.

Once again, syntactic and semantic local contexts are not discrete variables. However, this approach must also be contrasted with global semantic and pragmatic contexts, since conceptual knowledge as reflected in the text is not always reliable. This means that texts do not reflect perfectly designed conceptual networks. For instance, in hyponymic term-external semantic contexts (*x such as y, y and other x, x is a type of y,* etc.), authors do not always choose the direct parent of a concept. Many times, they will use a grandparent (WORK > GROIN instead of COASTAL STRUCTURE > GROIN) or will even create an *ad hoc* category (OBSTACLE TO FLOW > GROIN instead of COASTAL STRUCTURE > GROIN; León-Araúz and Reimerink, 2016).

Furthermore, the existence of multiple hyperonyms can indicate two types of multidimensionality: intracategorial and inter-categorial multidimensionality. In intracategorial multidimensionality, hyperonyms point to the same concept but highlight different dimensions or different levels of granularity. However, in intercategorial multidimensionality, hyperonyms point to a paradigm change, which makes the different facets incompatible.

One example is FOREST, which is found in local contexts as a type of ecosystem or as a type of renewable resource. This means that the concept is viewed as a type of one hyperonym to the exclusion of the others. This evidently affects the way in which the concept relates to other concepts. For instance, when a forest is viewed as a renewable resource, it is more closely related to concepts such as solar energy and biofuel, whereas when viewed as an ecosystem, wetlands and lakes are its closest concepts. This necessarily has an impact on knowledge and context modeling. Contrasting these results with a global approach (see section "Global Contexts") and analyzing lexical cohesion in the whole text where these structures occur can result in a reliable reconstruction of a text-driven conceptual system.

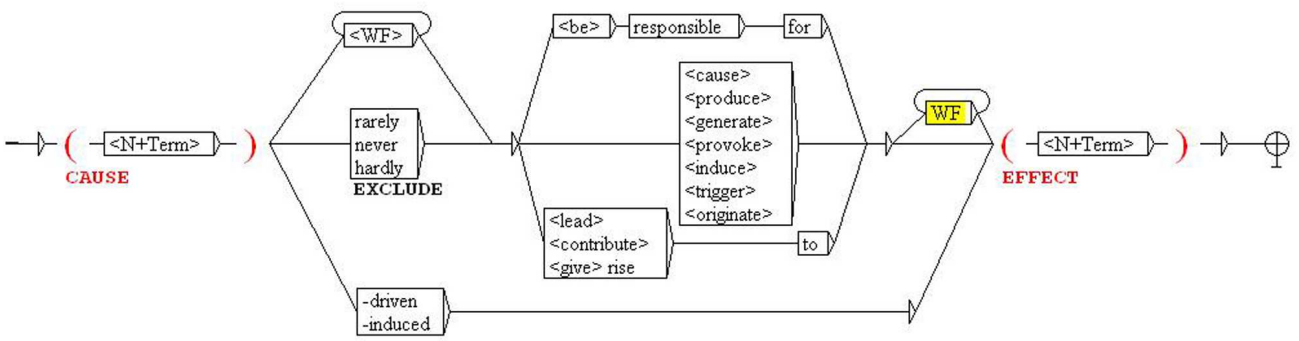

FIGURE 3 | Local grammar for causal relation extraction.

Local Pragmatic Contexts

Local pragmatic contexts basically refer to parameters of terminological variation and culturemes. Although in Terminology, the initial goal was to have one linguistic designation for each concept for greater precision, it soon became obvious that in descriptive terminology work, this is not always the case. This occurs more frequently in standardization settings (e.g., institutional, legal, technical, etc.) where the objective is to harmonize terminologies for the sake of efficient unambiguous communication. However, in the same way as for general concepts, the same specialized concept can often have many linguistic designations depending on the context. Alternatively, the same linguistic designation can also refer to various concepts.

As in general language, it is possible to establish reasons for terminological variation based on user-based parameters of geographic, temporal or social variation or usage-based parameters of field, tenor, and channel (Gregory and Carroll, 1978). Nevertheless, these basic parameters only provide a very partial representation of a very complex situation, since there are other reasons for terminological variation that are often considerably more difficult to represent.

Freixa (2006, p. 52), for example, classifies the causes for terminological variation in the following categories: (1) dialectal, based on the origin of the authors; (2) functional, based on communicative registers; (3) discursive, based on the style of the authors; (4) interlinguistic, based on the contact between languages; and (5) cognitive, based on different conceptualizations. These are all pointers to different types of extra-textual contexts, which mainly stem from the author's identity, location, language, and way of thinking. According to Freixa (2002), cognitive term variants are not only formally different, but also semantically diverse, as they give a particular vision of the concept. They are thus the natural reflection of multidimensionality (Fernández-Silva et al., 2011). Very often, the choice of one term instead of another stems from different perspectives of reality. Nevertheless, there are certain types of variation that do not fall into any of these categories, such as morphological variants, orthographic variants, ellipted variants, abbreviations, graphical variation, variation by permutation, etc. (Bowker and Hawkins, 2006, p. 81). Their use in texts often seems to be random without responding to any pattern or regularity. Although initially, the existence of such variants may not seem to be a problem, reality is somewhat different. Since term variants are rarely interchangeable, it is not a question of merely adding more terms to the TKB. What is needed is more information in term entries so that users can know which term to select. In terminological resources, users are often confronted with a vast array of variants with no indication of how term variation arises or how their use may be constrained.

In fact, variants often have a communicative and/or cognitive motivation. Therefore, the use of one term or another may affect the semantics of a concept or the communicative situation in which the concept is activated. Based on this distinction, our experience in EcoLexicon and other foundational work on term variation, we propose that term entries should include the following extended classification of pragmatic markers. When building a multilingual TKB, these markers can also enhance interlingual correspondences, because users will be able to make a cognitively sound choice. Otherwise, translators may actually over-standardize, creating consistency in places where the use of variants was deliberate and well-reasoned (Bowker and Hawkins, 2006, p. 80).

(A) *Orthographic variants* (with no geographic origin, e.g., *aesthetics*, *esthetics*), which do not affect semantics or the communicative situation.

(B) *Diatopic variants*

- (i) *Orthographic variants* (e.g., *groyne*, *groin*), which do not affect semantics.
- (ii) *Dialectal variants* (e.g., *gasoline*, *petrol*), which may affect semantics if culture-bound factors highlight or suppress any of the semantic features.
- (iii) *Culture-specific variants* (e.g., *dry lake*, and *sabkha*), which affect both semantics and the communicative situation. When referring to a particular entity that, in a specific culture, adds more specific features, culture-specific terms are used as variants of the closer entity in the target culture.
- (iv) *Calques*, which may affect semantics and the communicative situation and are the result of an interlinguistic borrowing for different reasons, such as the impact of a certain language on a specialized domain or the inexistence of the entity or term in a particular language.

(C) *Short form variants*, which do not affect semantics but only the communicative situation.

- (i) *Abbreviation*
- (ii) *Acronym* (e.g., *laser*, *Light Amplification by Stimulated Emission of Radiation*).

(D) *Diaphasic variants*

- (i) *Scientific variants*, which do not affect semantics but only the communicative situation.
  - (a) *Scientific names* (e.g., *Dracaena draco*, *drago*), which refer to specialized nomenclatures and are especially useful in botany, zoology, chemistry, etc.
  - (b) *Expert neutral variants* (e.g., *Ocellaris clownfish*, *Amphiprion ocellaris*), which would be the default term choice in a specialized scenario.
  - (c) *Jargon*. Sometimes experts have their own informal way to refer to specialized concepts (e.g., in medicine, *lap-appy* would correspond to *laparoscopic appendectomy*, but no lay user would use this term).
  - (d) *Formulas* (e.g., H_2O, *water*; $CaCO_3$, *pearl*), which do not affect semantics but only the communicative situation. H_2O and *water* do not change in meaning but the formula imposes certain constraints on text senders and receivers.
  - (e) *Symbols* (e.g., *€*, *euro*).
- (ii) *Informal variants*, which do not necessarily affect semantics but especially the communicative situation.

(a) *Lay user variants* (e.g., *Dragon tree, drago*), which would be the default term choice in non-specialized scenarios.
(b) *Colloquial variant* (e.g., *hydraulic fracturing*, *fracking*).
(c) *Generic variants* (e.g., *sea*, *ocean*; *weathering*, *erosion*). Very informal variants can activate terms pointing to different levels of conceptual granularity and thus affecting semantics.

(iii) *Domain-specific variants* (e.g., *mud*, *sludge*), which may affect semantics and/or the communicative situation when term preferences change across specialized domains.

(E) *Dimensional variants* (e.g., *Gutenberg's discontinuity*, *core-mantle boundary*), which are usually multi-word terms and affect semantics, since they convey different dimensions of the same concept (*Gutenberg's discontinuity* activates the person who first named it and *core-mantle boundary* the two areas that it separates).
(F) *Metonymic variants* (e.g., *water*, *sea*), which affect semantics because the metonymic variant designates the parts or material that the concept is made of.
(G) *Diachronic variants*.
(H) *Non-recommended variants* (e.g., in medicine, *mental retardation* now has negative connotations and has been substituted by *intellectual disability*).
(I) *Morpho-syntactic variants* (e.g., *wave action*, *the action of the waves*), which do not affect semantics but depend on collocates, term selection preferences, and the communicative situation.

The nature and scope of these variants are very diverse. Furthermore terms can activate more than one type of variant, which might make term choice more difficult. For example, H_2O and/or *water* are domain-based variants since the first is more frequently used in chemistry and water treatment domains than in hydrology or geology. However, their use also depends on the communicative situation and the knowledge level of the speaker and receiver. In the same way, *lap-appy* could be classified as jargon as well as a short form. On the other hand, the same type of variant can also be expressed by more than one term. Diaphasic variants, in particular, form a continuum ranging from more formal to informal (e.g., *thermal low pressure system*, *thermal low*, *thermal trough,* and *heat low*). The same happens within expert variants, which can be graded on a scale of frequency or acceptance. For instance, coastal defense and coastal management are both expert variants, but, coastal management is the preferred term.

Moreover, in those cases where the same concept can be designated by different dialectal variants stemming from the geographic origin of the writer, this can also mean that, conversely, the same term can be used to designate different entities in different cultures. For instance, *pier* (a structure built on posts extending from land out over water, used as a landing place for ships, an entertainment area, or a promenade area) is often designated as *jetty* in the Great Lakes. In contrast, a jetty is most often a structure designed to prevent the shoaling of a channel and is not considered a recreational area. However, in British English, *jetty* is the synonym of a *wharf*, whereas, in American English, *pier* may also be a synonym of *dock*. Nonetheless, in British English a *dock* is the area of water used for loading or unloading cargo in a harbor, which in American English is called a *port*.

Geographical variation in this category domain can often be conceptually motivated and mainly based on the dimensions of location and function. For instance, a dike may be called a *levee* when it is located on a river, whereas a breakwater may be called a *mole* when it is covered by a roadway. On the contrary, when a breakwater serves as a pier, it is called a *quay* in British English and a *wharf* in American English. Needless to say, when the knowledge base includes a conceptual representation or ontology, important design decisions must be taken. A base concept must be chosen (e.g., PIER) and be specified to accommodate the references of these variants (e.g., WORKING_PIER, PLEASURE_PIER, FISHING_PIER, etc.).

Local pragmatic contexts are thus reflected in terms and multiword expressions that are pointers to larger (global) situational, linguistic, and cultural contexts. Therefore, local and global pragmatic contexts constrain each other. Local contexts point to global contexts by constraining all possible situations (i.e., geographic, communicative, cognitive), and global contexts drive the choice of one variant over the rest.

Consequently, term variation should not be regarded as a linguistic phenomenon isolated from conceptual and cultural representations since it is one of the manifestations of the dynamicity of categorization and expression of specialized knowledge (Fernández-Silva et al., 2014).

Global Contexts

Contexts can also be global with a wider scope. The scope of such contexts can be a whole document, a communicative situation (e.g., formal vs. informal), a subject domain (e.g., Geology, Meteorology, etc.), or an entire language-culture.

Global contexts affect the underlying design of the data fields of a TKB since they are too large to be included in a term entry unless it is in the form of syntactic, semantic, or pragmatic markers, which would be more suitable for local contexts. They can also be analyzed with a view to tagging and classifying corpora in macro- and micro-structural terms.

The macro- and micro-structure of a text, or even a set of texts where intertextuality plays a role in understanding a specialized domain, provide a larger context to be analyzed with regards to grammatical and lexical cohesion (syntactic and semantic context). When global contexts are extra-textual, they are global pragmatic contexts characterized by different combinations of authorship, readership, function, domain, culture, etc.

Global Syntactic Contexts

When the document is used as the context, global syntactic contexts consist of the means of grammatical cohesion that tie the text together. These include endophoric reference (anaphora, cataphora), substitution, and ellipsis, as well as other grammatical cohesive markers that connect the different sentences of a text in a logical manner, such as, *however*, *on the other hand*, *consequently*,

etc. Such contexts could presumably also refer to the use of verb tenses throughout a discourse. For example, the typical Introduction, Methodology, Results, and Discussion (IMRAD) format of research articles is reflected in the verb tenses used in each section.

These verb tenses set the scene for the description of a research study and the presentation of results. In English, for example, the introduction is generally in the present tense when the author is describing the cause or (problematic) situation that produced the need for the research and the review of the literature on the topic. However, the past tense is used when referring to how the experiment was carried out. The present tense is again used when the author outlines the sections of the article. This use of verb tenses can differ when the article is written in a different language. For instance, in Spanish, there is a greater tendency to use the present and future tenses as though the research being described were being carried out in the paper itself.

Although it might seem that the use of tenses and syntax in general has little impact on terminology structure and selection, this is not the case. As observed by Gotti (2003), the fact that specialized discourse is characterized by elementary surface structures and relatively simple syntax allow the author a certain license to use complex and long pre-modified MWEs, which leads to a far longer sentence length. This means that the use of more or less complex nominal compounds is in direct relation to the relative simplicity of the syntactic structures in the text.

Grammatical cohesion in scientific discourse is often domain-independent, but still specialized. The same happens with the *transdisciplinary scientific lexicon* (Drouin, 2010), which includes abstract verbs (*to think*, *to consider*), abstract nouns (*idea*, *factor*, *relation*, *hypothesis*, *data*, *approach*), and collocations (*to conduct an analysis*) that refer to the description of scientific activities and reasoning but do not point to domain concepts. Thus, the study of global syntactic contexts can also have important computational applications, such as term extraction and coreference resolution.

Global Semantic Contexts

Global semantic contexts are in turn reflected in the lexical cohesion of texts (Halliday and Hasan, 1976; Morris and Hirst, 1991). Lexical cohesion is based on the meaning relations between words in a text. Such relations are paradigmatic and link two words having a common component from the viewpoint of their meaning.

Apart from repetition, lexical cohesion is most frequently achieved by using synonyms and hyperonyms, which requires a certain previous knowledge of the domain. In this sense, the description of local semantic and pragmatic contexts in TKBs ensures lexical cohesion when TKBs are used for text production tasks. In fact, it has been shown that scientific journal articles and popularized accounts of the same research do not employ the same cohesive patterns (Myers, 1991). According to Myers (1991, p. 5), the readers of scientific texts must have previous knowledge of lexical relations to see the implicit cohesion of the text, while readers of popularizations must see the explicitly marked cohesive relations to infer lexical relations, and to link the semantic field of the specialized domain to those of everyday life.

Thus, the analysis of lexical cohesive devices, which is hardly a trivial task, has also been approached from a computational perspective. In distributional approaches, synonyms, hyperonyms, antonyms, etc., are typically calculated by means of context vectors for each word, grouping together words that appear in the same contexts. In Ellman and Tait's (1998) framework, a single instance of lexical cohesion is a lexical link, whereas a sequence of links is a lexical chain. Such chains can also be formed by relations or bonds between sentences that are related by two or three links (Hoey, 1991).

Lexical chains are identified by using relationships between word senses. Nevertheless, in order to build lexical chains, it is necessary to know word senses and semantic relations between words. A lexical chain for a text contains a subset of the words (word senses) in the text, and are semantically related. Although the length of such chains may cover a larger or smaller portion of the text, in this case, we are referring to those that cover the whole document. Evidently, the number of words and the number of semantic relations between words can be different for each lexical chain. According to Ercan and Cicekli (2007), the coverage and size of a lexical chain can indicate how well the lexical chain represents the semantic content of the text. Lexical chains are evidently meaning-based but can also be derived from collocational frequencies. For example, Phillips (1989, p. 51) states that the collocation between *electric* and *charge* is also linked to the patterns in the text between their collocations (e.g., *charge* collocates with *distribution*, *density*, *point*, and *uniform*; *electric* collocates with *dipole*). Bondi (2010, p. 4) affirms that this network of semantic relations identifies the 'aboutness' of a text, and is a marker of text content.

Global Pragmatic Contexts

Global pragmatic contexts are the most complex form of context to be systematized and should thus be represented in a TKB, since they involve different interrelated variables.

Pragmatics is at the core of the dynamics of both terms and concepts (León-Araúz and Faber, 2014), since changes in conceptualization and in the lexicon are clearly not independent of each other but interact in a number of unforeseeable ways (Cimiano et al., 2010).

Generally speaking, pragmatics focuses on the effect of context on communicative behavior as well as on how inferences are made by the receiver (Faber, 2012). Crucial pragmatic dimensions in specialized communication contexts include (1) the beliefs and expectations of the text sender; (2) the knowledge shared by the text sender and text receivers; (3) the communicative objectives of the oral or written text stemming from the interaction of the participants; and (4) the factors that cause receivers to interpret the text in a certain way (Faber and San Martín, 2012, p. 178).

Strictly speaking, by its very definition, any type of pragmatic context is global. As previously mentioned, even local pragmatic contexts, as reflected in term variants or culturemes, are markers that point to larger communicative and cultural situations, which have an impact on conceptualizations in a given language-culture. Precisely for that reason, the description of entities is necessarily constrained by contextual

variation across communicative situations, cultures, and disciplines, as well as the fuzzy category boundaries that they establish.

For example, texts with a high term density (percentage of specialized knowledge units) are written by experts who wish to transmit knowledge to other experts in the same domain. Texts written for semi-experts or for non-experts have a correspondingly lower term density, although more term variants tend to be employed for the sake of transparency.

In specialized communication, genre and register are important concepts even though their definitions often seem to confusingly run together. However, following Lee (2001, pp. 46–47), we use register to refer to lexical-grammatical and semantic discourse patterns associated with situations, whereas genre is used to refer to the membership of a text in culturally recognizable categories, which may invoke more than one register. As such, genre is a socio-pragmatic phenomenon. According to Unger (2002, p. 2), a socio-pragmatic phenomenon is a set of shared assumptions that governs the communicative behavior of members of this group. It also relates communicative behavior to the structure of cultural institutions.

Hoffmann (1985, 1990) states that the purpose of a text depends on the context in which the text was created. In this sense, a text is both an instrument and a result that comes into being because of the specific productive activity (Hoffmann, 1985, p. 233). Similarly, Roelcke (1999, p. 42) underlines the importance of the specialized text regarded as a whole, and observes that the context of language usage also goes hand in hand with an increasing specialization of scientific and professional fields.

Although a definitive inventory and classification of specialized language genres and registers does not as yet exist, specialized language genres would doubtlessly be linked to specialized knowledge activities and text function within the context of a specialized knowledge field (cf. Hoffmann, 1985). Registers would presumably be subdivided primarily according to levels of formality. These formality levels would be constrained by parameters inherent in the context of specialized communication.

However, in TKBs, communicative context should not only be codified as a local pragmatic marker in term entries, especially when the aim of querying a TKB is multilingual communication. The reason for this is the fact that register-based variants in different languages do not necessarily establish 1:1 correspondences. This means that if a concept is designated by an informal term variant, it should not always be translated by its informal counterpart in another language and vice versa, because pragmatic conventions can also change from culture to culture. For instance, in an English doctor–patient communicative act, doctors tend to use more informal variants than in a similar situation in Spain. Even if a term-pair such as *intestinos* and *intestines* are full equivalents, *bowels* would be more appropriate in an English situation.

Nevertheless, the influence of culture is reflected in specialized domains in much more complex ways than it is in culture-specific terms or register-based differences (Faber and León-Araúz, 2014). They also may affect conceptual structures. For instance, one might think that natural landforms are more or less the same all over the world, but the truth is that there is a great deal of plasticity in how language models the earth and what is considered to be the essence of its features (Burenhult and Levinson, 2008, p. 148). Until recently, it was believed that entities such as mountain and river were candidates for universals (Smith and Mark, 2001). However, research in cognitive ethnophysiography has found that this is not the case. Apart from the problem of establishing interlinguistic correspondences, this also makes it hard to agree on how concepts are classified in the same language.

For example, the diversity of wetlands is an obstacle to arriving at a consensus in regards to their classification. One of the most widely used classifications was created by Cowardin et al. (1979), who divided wetlands into marine, estuarine, riverine, lacustrine, and palustrine environments. Nevertheless, this classification was eventually found to be too restrictive, and a more comprehensive categorization was required. The Ramsar classification system for wetland types (1996) thus proposed new categories to cover all types of wetlands in the world: marine/coastal wetlands, inland wetlands, human-made wetlands. In turn, the Canadian national wetlands working group (1997) established five classes: bog, fen, marsh, swamp, and shallow water.

However, labeling categories in terms of basic level concepts (Rosch, 1978) can be confusing, because they are highly localized. For instance, bogs or fens are usually grouped together and referred to as *mires* in Europe, but not in America. Marshes in Europe are often called *reed swamps*, but swamps in America are not dominated by reeds but rather by trees. *Carr* is the northern European term for the Southeast American *wooden swamp*, which in the United Kingdom is also called *wet woodland*. There are also specific types of wetlands that only predominate in certain geographic areas that are not lexicalized in all cultures, such as the Australian *billabong*, the African *dambo,* or the Canadian *muskeg*. In these cases, the local terms are only borrowed when describing these particular wetlands. Thus, when one of these terms is activated in a text, the location-related category features of the concept are constrained.

Multidimensionality is also found in discipline-based contexts. In Terminology, multidimensionality is often regarded as a way of enriching traditional static representations, enhancing knowledge acquisition through different points of view in the same semantic network or conceptual system. However, it can also produce an excessive information load. This is the case of certain general top-level concepts such as water (**Figure 4**), which is a classic example of information overload in EcoLexicon (León-Araúz et al., 2012, 2013; Faber et al., 2014a).

Water certainly holds different relations with a myriad of different concepts. However, EcoLexicon users would not acquire any meaningful knowledge if all dimensions of water were shown in the same network. Moreover, water rarely, if ever, activates those concepts at the same time, since this would evoke completely different and incompatible scenarios (León-Araúz et al., 2013). In this sense, although it is true that concepts cannot be activated in isolation, they can also retain sufficient autonomy so that the activation of one does

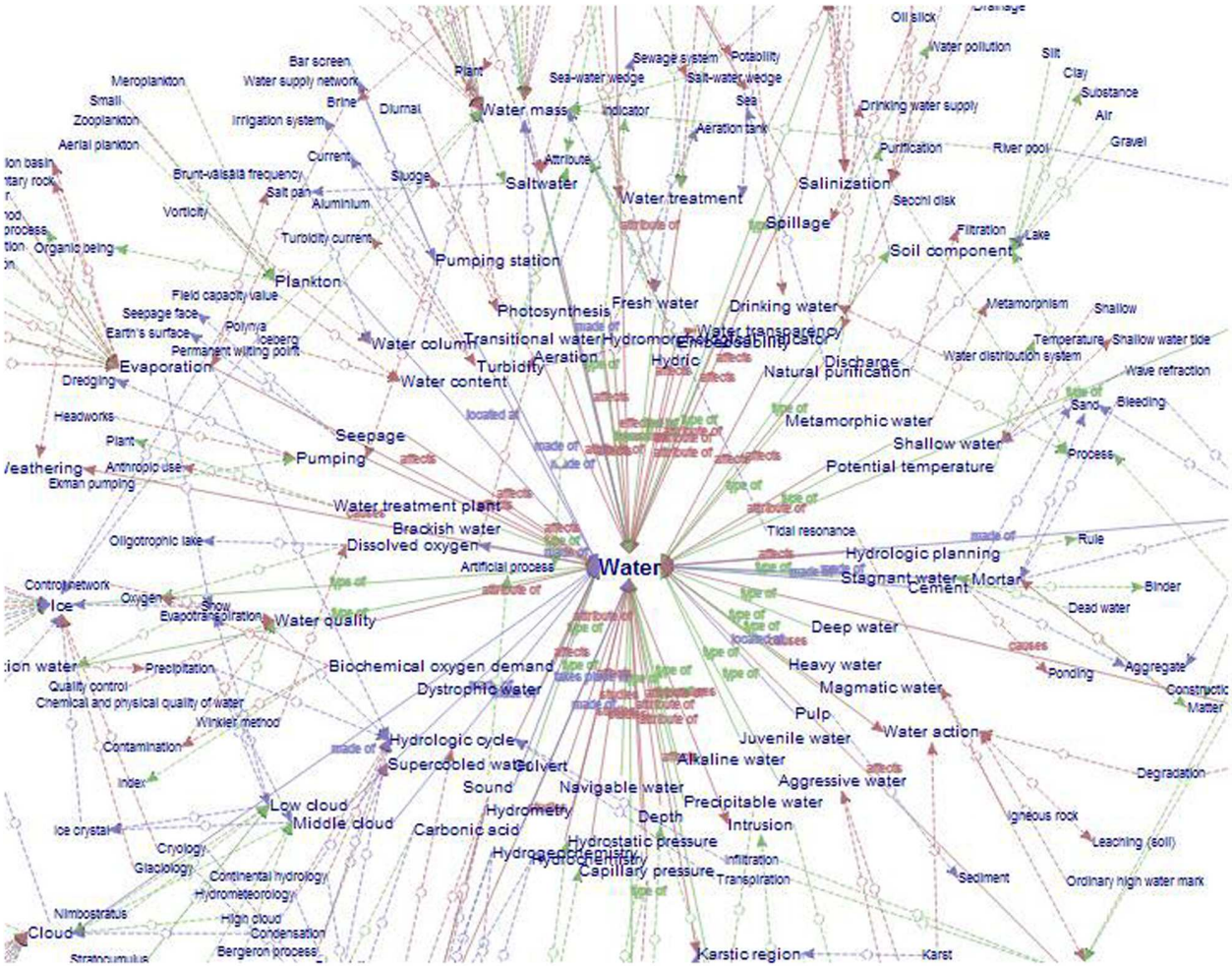

FIGURE 4 | Information overload in EcoLexicon.

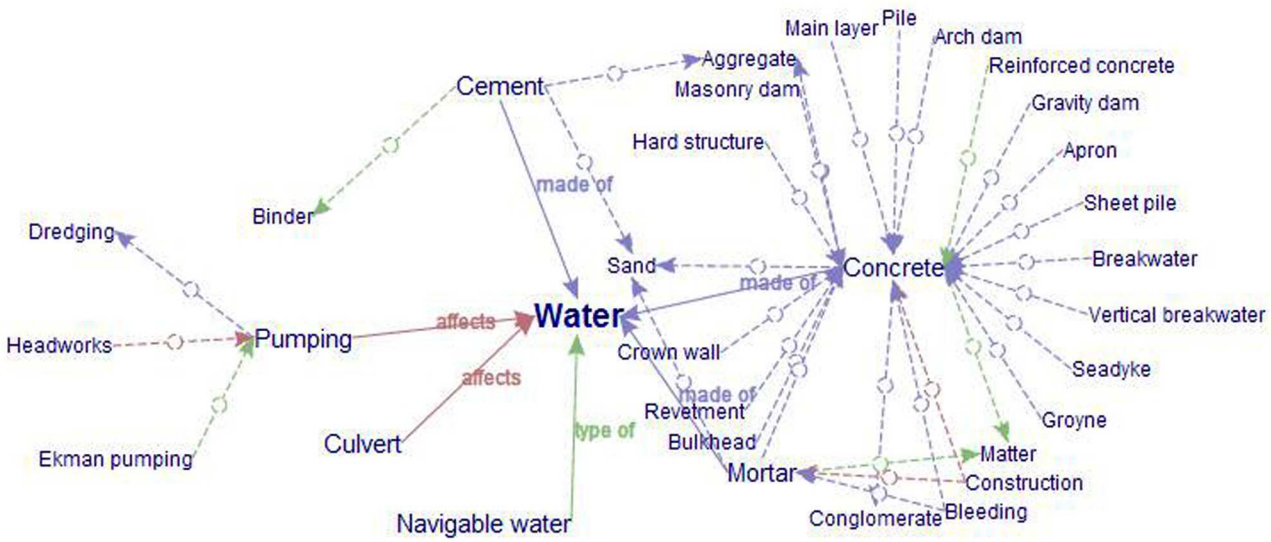

FIGURE 5 | WATER as recontextualized in Civil Engineering.

not necessarily entail the activation of the rest (Langacker, 1987, p. 162). Their activation should thus be domain-dependent.

According to Picht and Draskau (1985, p. 48), multidimensionality depends on the classifier as well as the different knowledge sources that may reflect different criteria when organizing the same domain. In conceptual modeling, facets and contexts can be established according to different criteria. However, in EcoLexicon, a discipline-oriented approach was found to be the most appropriate, since concepts may have different roles and degrees of prominence in the different disciplines that constitute the environmental sciences.

As opposed to formal approaches where concepts are ascribed to particular categories on the basis of a set of necessary and sufficient features, semantic networks in EcoLexicon take the form of a set of conceptual relations that might be highlighted or suppressed, depending on pragmatic factors. We agree with Michalski (1991) when he states that the context of a concept is the set of concepts relevant to its intended meaning.

The environmental domain was thus divided into a set of domain-based contexts (e.g., hydrology, geology, oceanography, civil engineering, etc.) and the relational power of concepts was constrained accordingly. This is done by assigning each conceptual proposition to one or more contextual domains based on a previously domain-based classified corpus. For example, the proposition CONCRETE *made_of* WATER only appears relevant in Civil Engineering texts, but not in a geological context. Thus, when constraints are applied, the network of WATER within the civil engineering sub-domain is recontextualized and becomes more meaningful (**Figure 5**).

Recontextualization is in line with Cruse's (2002) approach to meaning (i.e., ways of seeing, microsenses, or context modulation) or Sperber and Wilson's (1986) Relevance Theory, since semantic networks are dynamically built according to context salience. Thus, concepts themselves can also have their own situated nature. In this sense, Barsalou (2009, p. 1283) states that a concept produces a wide variety of situated conceptualizations that support goal achievement in specific contexts. In a similar way, this would be in consonance with semantic priming, which according to McNamara (2005) can be influenced by the context created by the types of semantic relations present in a test list.

CONCLUSION AND FUTURE WORK

In this paper, we have proposed a taxonomy of context primarily based on scope (local and global) and further divided into syntactic, semantic, and pragmatic facets for TKB design. Although context is a controversial notion interpreted and represented as needed in each field, we believe that for specialized knowledge representation in terminological resources, context should be much more than a textual excerpt.

Context modeling formally describes aspects of the linguistic, physical, and social world around us for purposes of understanding and communication. In this regard, it is necessary to determine what aspects to include and exclude from the model, and at what level of detail to model each of them.

Ideally, context specification and representation in specialized knowledge resources is conducive to the formulation of a common structure applicable to and valid for different languages and cultures based on a representational framework that allows for correspondences at different levels as well as for the inclusion of the syntactic, semantic and pragmatic features upon which this correspondence is based.

In EcoLexicon, various context parameters have been explored. Their specification has materialized in various modules representing different contextual aspects, ranging from term variation, collocations or knowledge-rich contexts, to dynamic conceptual networks, flexible definitional templates, or conceptually enriched graphical resources. However, much still remains to be done, especially with regards to the interdependence between all modules and the transition from local to global constraints. In the future, users will be provided with different types of information selected according to context. For this to be possible, context in all of its facets must be accounted for in a systematic and principled way.

AUTHOR CONTRIBUTIONS

All authors listed, have made substantial, direct and intellectual contribution to the work, and approved it for publication.

ACKNOWLEDGMENTS

This research was carried out within the framework of project FF2014-52740-P, Cognitive and Neurological Bases for Terminology-enhanced Translation (CONTENT) funded by the Spanish Ministry of Economy and Competitiveness.

REFERENCES

Agirre, E., and Stevenson, M. (2007). "Knowledge sources for word sense disambiguation," in *Word Sense Disambiguation: Algorithms and Applications*, eds E. Agirre and P. Edmonds (New York, NY: Springer).

Akman, V. (2000). Rethinking context as a social construct. *J. Pragmat.* 32, 743–759. doi: 10.1016/S0378-2166(99)00067-3

Akman, V., and Bazzanella, C. (2003). The complexity of context: guest editors' introduction. *J. Pragmat.* 35, 321–329. doi: 10.1016/S0378-2166(02)00138-8

Akman, V., and Surav, M. (1997). The use of situation theory in context modeling. *Comput. Intell.* 13, 427–438. doi: 10.1111/0824-7935.00046

Aussenac-Gilles, N., and Jacques, M. P. (2006). "Designing and evaluating patterns for ontology enrichment from texts," in *International Conference on Knowledge Engineering and Knowledge Management EKAW 2006. Lecture Notes in Artificial Intelligence 4248*, eds S. Staad and V. Svatek (Berlin: Springer), 158–165.

Austin, J. L. (1962). *How to Do Things with Words*. Oxford: Oxford University Press.

Barrière, C. (2004). "Knowledge-rich contexts discovery," in *Proceedings of the 17th Conference of the Canadian Society for Computational Studies of Intelligence on Advances in Artificial Intelligence. Canadian AI 2004,* London.

Barsalou, L. W. (1983). Ad hoc categories. *Mem. Cogn.* 11, 211–227. doi: 10.3758/BF03196968

Barsalou, L. W. (1991). "Deriving categories to achieve goals," in *The Psychology of Learning and Motivation: Advances in Research and Theory*, Vol. 27, ed. G. H. Bower (San Diego, CA: Academic Press), 1–64.

Barsalou, L. W. (2009). Simulation, situated conceptualization, and prediction. *Philos. Trans. R. Soc. Lond. Biol. Sci.* 364, 1281–1289. doi: 10.1098/rstb.2008.0319

Bazire, M., and Brézillon, P. (2005). "Understanding context before using it," in *Modeling and Using Context*, eds A. Dey, B. Kokinov, D. Leake, and R. Turner (Berlin: Springer).

Bazzanella, C. (1998). "On context and dialogue," in *Dialogue in the Heart of Europe, Proceedings IADA (Prague 1996)*, eds S. Cmejrková, J. Hoffmannová, O. Müllerová, and J. Svetla (Tübingen: Niemeyer), 407–417.

Blommaert, J. (2005). *Discourse. A Critical Introduction.* Cambridge: Cambridge University Press.

Bondi, M. (2010). "Perspectives on keywords and keyness: an introduction," in *Keyness in Texts*, eds M. Bondi and M. Scott (Amsterdam: John Benjamins), 1–18.

Bowker, L. (1997). "Multidimensional classification of concepts and terms," in *Handbook of Terminology Management: Basic Aspects of Terminology Management*, eds S. E. Wright and G. Budin (Amsterdam: John Benjamins), 133–143.

Bowker, L. (2011). "Off the record and on the fly," in *Corpus-based Translation Studies: Research and Applications*, eds A. Kruger, K. Wallmach, and J. Munday (New York, NY: Continuum), 211–236.

Bowker, L., and Hawkins, S. (2006). Variation in the organization of medical terms. Exploring some motivations for term choice. *Terminology* 12, 79–110.

Brézillon, P. (2005). "Task-realization models in contextual graphs," in *Modeling and Using Context*, eds A. Dey, B. Kokinov, D. Leake, and R. Turner (Berlin: Springer).

Brown, J. S., Collins, A., and Duguid, P. (1989). Situated cognition and the culture of learning. *Educ. Res.* 18, 32–41. doi: 10.3102/0013189X018001032

Burenhult, N., and Levinson, S. C. (2008). Language and landscape: a crosslinguistic perspective. *Lang. Sci.* 30, 135–150. doi: 10.1016/j.langsci.2006.12.028

Cabré, M. T. (1999). *La Terminología: Representación y Comunicación. Elementos Para una Teoría de Base comunicativa y Otros Artículos*. Barcelona: IULA, Universidad Pompeu Fabra.

Cilibrasi, R. L., and Vitanyi, P. M. B. (2007). The google similarity distance. *IEEE Trans. Knowl. Data Eng.* 19, 370–383. doi: 10.1109/TKDE.2007.48

Cimiano, P., Montiel Ponsoda, E., Buitelaar, P., Espinoza, M., and Gómez Pérez, A. (2010). A note on ontology localization. *J. Appl. Ontol.* 0, 1–10. doi: 10.1186/s13567-015-0196-x

Clancey, W. J. (1994). "Situated cognition: how representations are created and given meaning," in *Lessons from Learning*, eds R. Lewis and P. Mendelsohn (Amsterdam: North-Holland), 231–242.

Clark, H. H. (1996). *Using Language*. Cambridge: Cambridge University Press.

Cowardin, L. M., Carter, V., Golet, F. C., and LaRoe, E. T. (1979). *Classification of Wetlands and Deepwater Habitats of the United States*. Washington, DC: U.S. Department of the Interior, Fish and Wildlife Service.

Croft, W., and Cruse, D. A. (2004). *Cognitive Linguistics*. Cambridge: Cambridge University Press.

Cruse, D. A. (2002). "Hyponymy and its varieties," in *The Semantics of Relationships*, eds R. Green, C. A. Bean, and S. H. Myaeng (Dordrecht: Kluwer), 3–21.

Dash, N. S. (2008). Context and contextual word meaning. *SKASE J. Theor. Linguist.* 5, 21–31.

Dey, A. K. (2001). Understanding and using context. *Pers. Ubiquitous Comput.* 5, 4–7. doi: 10.1007/s007790170019

Drouin, P. (2010). "Extracting a bilingual transdisciplinary scientific lexicon," in *Proceedings of the eLex 2009* on *E-Lexicography in the* 21st Century: New Challenges, New Applications, Louvain-la-Neuve, 43–53.

Dubuc, R., and Lauriston, A. (1997). "Terms and contexts," in *Handbook of Terminology Management*, eds S. E. Wright and G. Budin (Amsterdam: John Benjamins), 80–87.

Ellman, J., and Tait, J. (1998). "Meta searching the web using exemplar texts: initial results," in *Proceedings of the 20th Annual Colloquium on IR Research* (Glasgow: British Computer Society's Information Retrieval Specialist Group).

Elman, J. (2009). On the meaning words and dinosaur bones: lexical knowledge without a lexicon. *Cogn. Sci.* 33, 547–582. doi: 10.1111/j.1551-6709.2009.01023.x

Ercan, G., and Cicekli, I. (2007). Using lexical chains for keyword extraction. *Inform. Process. Manag.* 43, 1705–1714. doi: 10.1016/j.ipm.2007.01.015

Erman, B., and Warren, B. (2000). The idiom principle and the open-choice principle. *Text* 20, 29–62.

Ervin-Tripp, S. M. (1996). "Context in language," in *Social Interaction, Social Context and Language: Essays in Honor of Susan Ervin-Tripp*, eds D. I. Slobin, J. Gerhardt, A. Kyratzis, and J. Guo (Mahwah, NJ: Lawrence Erlbaum), 21–36.

Evans, V., and Green, M. (2006). *Cognitive Linguistics: An Introduction*. Edinburgh: Edinburgh University Press.

Faber, P. (2011). The dynamics of specialized knowledge representation: simulational reconstruction or the perception–action interface. *Terminology* 17, 9–29. doi: 10.1075/term.17.1.02fab

Faber, P. (ed). (2012). *A Cognitive Linguistics View of Terminology and Specialized Language*. Berlin: de Gruyter.

Faber, P. (2015). "Frames as a framework for terminology," in *Handbook of Terminology*, eds H. J. Kockaert and F. Steurs (Philadelphia: John Benjamins), 14–33.

Faber, P., and León-Araúz, P. (2014). "Specialized knowledge dynamics: from cognition to culture-bound terminology," in *Dynamics and Terminology. An Interdisciplinary Perspective on Monolingual and Multilingual Culture-Bound Communication, Terminology and Lexicography Research and Practice*, Vol. 16, eds R. Temmerman and M. Van Campenhoudt (Amsterdam: John Benjamins), 135–158.

Faber, P., León-Araúz, P., Prieto Velasco, J. A., and Reimerink, A. (2007). Linking images and words: the description of specialized concepts. *Int. J. Lexicogr.* 20, 39–65. doi: 10.1093/ijl/ecl038

Faber, P., León Araúz, P., and Reimerink, A. (2014a). "Representing environmental knowledge in ecoLexicon," in *Languages for Specific Purposes in the Digital Era. Educational Linguistics*, Vol. 19, eds E. Bárcena, T. Read, and J. Arús (Berlin: Springer), 267–301.

Faber, P., and Mairal Usón, R. (1999). *Constructing a Lexicon of English Verbs*. Berlin: Mouton de Gruyter.

Faber, P., Márquez Linares, C., and Vega Expósito, M. (2005). Framing terminology: a process-oriented approach. *Meta* 50. doi: 10.7202/019916ar

Faber, P., Montero Martínez, S., Castro Prieto, M. C., Senso Ruiz, J., Prieto Velasco, J. A., León-Araúz, P., et al. (2006). Process-oriented terminology management in the domain of Coastal Engineering. *Terminology* 12, 189–213.

Faber, P., and San Martín, A. (2012). "Specialized language pragmatics," in *A Cognitive Linguistics View of Terminology and Specialized Language*, ed. P. Faber (Berlin: De Gruyter Mouton), 177–203.

Faber, P., Verdejo, J., León-Araúz, P., Reimerink, A., and Guzman, G. (2014b). Neural substrates of specialized knowledge representation: an fMRI study. *Rev. Française Linguist. Appl.* 19, 15–34.

Fauconnier, G. (1985). *Mental Spaces*. Cambridge: Cambridge University Press.

Fauconnier, G. (1997). *Mappings in Thought and Language*. Cambridge: Cambridge University Press.

Fernández-Silva, S., Freixa, J., and Cabré, M. T. (2011). A proposed method for analyzing the dynamics of cognition through term variation. *Terminology* 17, 49–73. doi: 10.1075/term.17.1.04fer

Fernández-Silva, S., Freixa, J., and Cabré, M. T. (2014). "A proponed method for analysing the dynamics of naming from an intralingual and interlingual perspective," in *Dynamics and Terminology. An Interdisciplinary Perspective on Monolingual and Multilingual Culture-Bound Communication*, eds R. Temmerman and M. Van Campenhoudt (Philadelphia: John Benjamins), 183–211.

Firth, J. R. (1957). *Papers in Linguistics 1934-1951*. London: Oxford University Press.

Firth, J. R. (1968). "Linguistic analysis as a study of meaning," in *Selected Papers of J. R*, ed. F. R. Palmer (Firth: Longmans Linguistics Library), 12–26.

Flowerdew, J. (ed). (2014). *Discourse in Context*. London: Bloombury.

Freixa, J. (2002). *La Variació Terminològica: Anàlisi de la Variació Denominativa en Textos de Diferent Grau d'especialització de l'àrea de Medi Ambient*. Ph.D. thesis, Universitat Pompeu Fabra: Barcelona.

Freixa, J. (2006). Causes of denominative variation in terminology, A typology proposal. *Terminology* 12, 51–77. doi: 10.1075/term.12.1.04fre

Gabrielatos, C. (2007). Selecting query terms to build a specialised corpus from a restricted-access database. *ICAME J.* 31, 5–43.

Gadamer, H. G. (1995). *Truth and Method*. New York, NY: Continuum.

Goffman, E. (1974). *Frame Analysis*. New York, NY: Harper and Row.

Gotti, M. (2003). *Specialized Discourse. Linguistic Features and Changing Conventions*. Bern: Peter Lang.

Gregory, M., and Carroll, S. (1978). *Language and Situation: Language Varieties and their Social Contexts*. London: Routledge.

Grice, H. P. (1975). "Logic and conversation," in *Syntax and Semantics: Speech Acts*, Vol. 3, eds P. Cole and J. L. Morgan (New York, NY: Academic press), 41–58.

Gries, S., and Divjak, D. (2009). "Behavioral profiles: a corpus-based approach towards cognitive semantic analysis," in *New Directions in Cognitive Linguistics*, eds V. Evans and S. Pourcel (Philadelphia, PA: John Benjamins), 57–75.

Guha, R. V. (1991). *Contexts: A Formalization and Some Applications.* Ph.D. thesis, Stanford University, Stanford, CA.

Halliday, M. A. K. (1989). *Spoken and Written Language*. Oxford: Oxford University Press.

Halliday, M. A. K., and Hasan, R. (1976). *Cohesion in English*. London: Longman.

Harris, Z. (1985). "Distributional structure," in *The Philosophy of Linguistics*, ed. J. J. Katz (New York, NY: Oxford University Press), 26–47.

Harris, Z. (1988). *Language and Information*. New York, NY: Columbia University Press.

He, Q., Pei, J., Kifer, D., Mitra, P., and Giles, C. L. (2010). "Context-aware citation recommendation," in *Proceedings of WWW 10th Proceedings of the 19th International Conference on World Wide Web* (New York, NY: Association for Computing Machinery), 421–430.

Hoey, M. (1991). *Patterns of Lexis in Text*. Oxford: Oxford University Press.

Hoey, M. (2005). *Lexical Priming. A New Theory of Words and Language*. London: Routledge.

Hoffmann, L. (1985). *Kommunikationsmittel Fachsprache*. Berlin: Akademic.

Hoffmann, L. (1990). *Fachtexte und Fachtextsorten*. Leipzig: University of Leipzig.

House, J. (2006). Text and context in translation. *J. Pragma.* 38, 338–358. doi: 10.1016/j.pragma.2005.06.021

Jackendoff, R. (1997). *The Architecture of the Language Faculty*. Cambridge, MA: MIT Press.

Kageura, K. (1997). "Multifaceted/multidimensional concept systems," in *Handbook of Terminology Management: Basic Aspects of Terminology Management*, eds S. E. Wright and G. Budin (Amsterdam: John Benjamins), 119–132.

Langacker, R. W. (1987). *Foundations of Cognitive Grammar: Theoretical Prerequisites*, Vol. 1. Stanford: Stanford University Press.

Lee, D. (2001). Genres, registers, text types, domains, and styles: clarifying the concepts and navigating a path through the BNC jungle. *Lang. Learn. Technol.* 5, 37–72.

Leech, G. (1981). *Semantics: The Study of Meaning*. Harmondsworth: Penguin.

Lenat, D. B. (1995). CYC: a large-scale investment in knowledge infrastructure. *Commun. ACM* 38, 33–38. doi: 10.1145/219717.219745

León-Araúz, P., and Faber, P. (2014). "Context and terminology in the multilingual semantic web," in *Towards the Multilingual Semantic Web. Principles, Methods and Applications*, eds P. Buitelaar and P. Cimiano (Berlin: Springer).

León-Araúz, P., Faber, P., and Montero Martínez, S. (2012). "Specialized language semantics," in *A Cognitive Linguistics View of Terminology and Specialized Language*, Vol. 20, ed. P. Faber (Berlin: De Gruyter Mouton), 95–175.

León-Araúz, P., and Reimerink, A. (2016). "Multidimensional categorization in corpus-based hyponymyc structures," in *Corpus-based Studies on Language Varieties*, eds F. Alonso Almeida, L. Cruz García, and V. González Ruiz (Bern: Peter Lang).

León-Araúz, P., Reimerink, A., and Aragón, A. G. (2013). Dynamism and context in specialized knowledge. *Terminology* 19, 31–61. doi: 10.1075/term.19.1.02leo

L'Homme, M. C. (2003). Capturing the lexical structure in special subject fields with verbs and verbal derivatives: a model for specialised lexicography. *Int. J. Lexicogr.* 16, 403–422. doi: 10.1093/ijl/16.4.403

Louw, W. E. (1993). "Irony in the text or insincerity in the writer? The diagnostic potential of semantic prosodies," in *Text and Technology: In Honour of John Sinclair*, eds M. Baker, G. Francis, and E. Tognini-Bonelli (Amsterdam: John Benjamins), 157–176.

Maguire, P., Edward, J. W., and Gert, S. (2010). A corpus study of semantic patterns in compounding. *Corpus Linguist. Linguist. Theory* 6, 49–73. doi: 10.1515/cllt.2010.003

McCarthy, J. (1987). Generality in artificial intelligence. *Commun. ACM* 30, 1030–1035. doi: 10.1145/33447.33448

McCarthy, J. (1993). "Notes on formalizing context," in *Proceedings of the 13 International Joint Conference on Artificial Intelligence*, Chambéry, 555–560.

McNamara, T. (2005). *Semantic Priming: Perspectives from Memory and Word Recognition. Essays in Cognitive Psychology*. New York, NY: Psychology Press.

Meyer, I. (2001). "Extracting knowledge-rich contexts for terminography: a conceptual and methodogical framework," in *Recent Advances in Computational Terminology*, eds D. Bourigault, C. Jacquemin, and M. C. L'Homme (Amsterdam: John Benjamins), 279–302.

Meyer, I., Skuce, D., Bowker, L., and Eck, K. (1992). "Towards a new generation of terminological resources: an experiment in building a terminological knowledge base," in *Proceedings of the 14th International Conference on Computational Linguistics (COLING 92)*, Nantes, 956–960.

Michalski, R. S. (1991). *Concepts as Flexible and Context-Dependent Sets: The Two-Tiered View*. Fairfax: George Mason University.

Miecznikowski, J., and Bazzanella, C. (2007). "The attenuating conditional: context, appropriateness and interaction," in *Context and Appropriateness. Micro meets Macro*, ed. A. Fetzer (Amsterdam: John Benjamins), 203–233.

Mihalcea, R. (2007). "Using Wikipedia for automatic word sense disambiguation," in *Proceedings of NAACL HLT* (Rochester, NY: Association for Computational Linguistics), 196–203.

Miller, G. A., and Charles, W. G. (1991). Contextual correlates of semantic similarity. *Lang. Cogn. Process.* 6, 1–28.

Mitkov, R. (1998). The latest in anaphora resolution: going robust, knowledge-poor and multilingual. *Procesamiento Leng. Nat.* 23, 1–7.

Mitkov, R. (2002). *Anaphora Resolution*. London: Longman.

Mollin, S. (2009). Combining corpus linguistic and psychological data on word co-ocurrences: corpus collocates versus word associations. *Corpus Linguist. Linguist. Theory* 5, 175–200. doi: 10.1515/CLLT.2009.008

Morris, J., and Hirst, G. (1991). Lexical cohesion computed by thesaural relations as an indicator of the structure of text. *Comput. Linguist.* 17, 21–48.

Murphy, G. L. (1988). Comprehending complex concepts. *Cogn. Sci.* 12, 529–562. doi: 10.1207/s15516709cog1204_2

Myers, G. (1991). Lexical cohesion and specialized knowledge in science and popular science texts. *Discourse Processes* 14, 1–26. doi: 10.1080/01638539109544772

Pearson, J. (1998). *Terms in Context*. Amsterdam: John Benjamins.

Phillips, M. (1989). *Lexical Structure of Text. In Discourse Analysis Monographs 12*. Birmingham: University of Birmingham.

Picht, H., and Draskau, J. (1985). *Terminology: An Introduction*. Guildford: University of Surrey.

Pustejovksy, J. (1995). *The Generative Lexicon*. Cambridge, MA: MIT Press.

Reimerink, A., García de Quesada, M., and Montero-Martínez, S. (2010). Contextual information in terminological knowledge bases: A multimodal approach. *J. Pragmat.* 42, 1928–1950. doi: 10.1016/j.pragma.2009.12.008

Roelcke, T. (1999). *Fachsprachen*. Berlin: Erich Schmidt.

Rogers, M. (2004). Multidimensionality in concepts systems: a bilingual textual perspective. *Terminology* 10, 215–240. doi: 10.1075/term.10.2.04rog

Rosch, E. (1978). "Principles of categorization," in *Cognition and Categorization*, eds E. Rosch and B. L. Lloyd (Hillsdale, NJ: Lawrence Erlbaum), 27–47.

Rumshisky, A. (2008). Resolving polysemy in verbs: contextualized distributional approach to argument semantics. *Riv. Linguist.* 20, 215–240.

Sager, J. C. (1990). *A Practical Course in Terminology Processing*. Amsterdam: John Benjamins, 197–205.

Searle, J. R. (1969). *Speech Acts: An Essay in the Philosophy of Language*. Cambridge: University Press.

Sierra, G., Alarcón, R., Aguilar, C., and Bach, C. (2008). Definitional verbal patterns for semantic relation extraction. *Terminology* 14, 74–98. doi: 10.1075/term.14.1.05sie

Smith, B., and Mark, D. (2001). Geographic categories: an ontological investigation. *Int. J. Geogr. Inform. Sci.* 15, 591–612. doi: 10.1080/13658810110061199

Soricut, R., and Marcu, D. (2003). "Sentence level discourse parsing using syntactic and lexical information," in *Proceedings of the Human Language Technology and North American Association for Computational Linguistics Conference (HLT/NAACL)*, Edmonton, AB.

Sperber, D., and Wilson, D. (1986). *Relevance: Communication and Cognition*, 2nd Edn. Oxford: Blackwell.

Sperber, D., and Wilson, D. (1990). Spontaneous deduction and mutual knowledge. *Behav. Brain Sci.* 13, 179–184. doi: 10.1017/S0140525X00078237

Stubbs, M. (2002). Two quantitative methods of studying phraseology in English. *Int. J. Corpus Linguist.* 7, 215–244. doi: 10.1075/ijcl.7.2.04stu

Tesnière, L. (1959). *Éléments de Syntaxe Structurale*. Klincksieck: Paris.

Thagard, P. (1989). Explanatory coherence. *Behav. Brain Sci.* 12, 435–467. doi: 10.1017/S0140525X00057046

Thagard, P., and Verbeurgt, K. (1998). Coherence as constraint satisfaction. *Cogn. Sci.* 22, 1–24. doi: 10.1207/s15516709cog2201_1

Unger, C. (2002). "Cognitive-pragmatic explanations of socio-pragmatic phenomena: the case of genre," in *Proceedings of the I Symposium Conference Paper Read at the EPICS*, Seville. Available at: http://cogprints.org/5439/

Van Dijk, T. A. (2005). *Discurso, Conocimiento e Ideología. Cuadernos de Información y Comunicación*. Madrid: Universidad Complutense de Madrid.

Wisniewski, E. J. (1997). When concepts combine. *Psychon. Bull. Rev.* 4, 167–183. doi: 10.3758/BF03209392

Wright, S. E. (1997). "Representation of concept systems," in *Handbook of Terminology Management*, eds S. E. Wright and G. Budin (Amsterdam: John Benjamins), 89–97.

Wüster, E. (1979). *Einführung in die Allgemeine Terminologielehre und Terminologische Lexikographie. Teil* 1-2. Berlin: Springer.

Yeh, W., and Barsalou, L. W. (2006). The situated nature of concepts. *Am. J. Psychol.* 119, 349–384. doi: 10.2307/20445349

Yus, F. (2006). "Relevance theory," in *Encyclopedia of Language and Linguistics*, 2nd Edn, Vol. 10, ed. K. Brown (Amsterdam: Elsevier), 512–518.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Faber and León-Araúz. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.