

Article

Are We Talking about the Same Thing? Modeling Semantic Similarity between Common and Specialized Lexica in WordNet

Chiara Barbero *  and Raquel Amaro 

Faculty of Social Sciences and Humanities, Linguistics Research Centre, NOVA University Lisbon, Avenida de Berna, 26-C, 1069-061 Lisbon, Portugal; raquelamaro@fcs.unl.pt

* Correspondence: chiarabarbero@fcs.unl.pt

Abstract: Specialized languages can activate different sets of semantic features when compared to general language or express concepts through different words according to the domain. The specialized lexicon, i.e., lexical units that denote more specific concepts and knowledge emerging from specific domains, however, co-exists with the common lexicon, i.e., the set of lexical units that denote concepts and knowledge shared by the average speakers, regardless of their specific training or expertise. Communication between specialists and non-specialists can show a big gap between language(s), and therefore lexical units, used by the two groups. However, quite often, semantic and conceptual overlapping between specialized and common lexical units occurs and, in many cases, the specialized and common units refer to close concepts or even point to the same reality. Considering the modeling of meaning in functional lexical resources, this paper puts forth a solution that links common and specialized lexica within the WordNet model framework. We propose a new relation expressing semantic proximity between common and specialized units and define the conditions for its establishment. Besides contributing to the observation and understanding of the process of knowledge specialization and its reflex on the lexicon, the proposed relation allows for the integration of specialized and non-specialized lexicons into a single database, contributing directly to improving communication in specialist/non-specialist contexts, such as teaching–learning situations or health professional–patient interactions, among many others, where code-switching is frequent and necessary.



Citation: Barbero, Chiara, and Raquel Amaro. 2024. Are We Talking about the Same Thing? Modeling Semantic Similarity between Common and Specialized Lexica in WordNet.

Languages 9: 89. <https://doi.org/10.3390/languages9030089>

Academic Editor: John Lipski

Received: 15 January 2024

Revised: 16 February 2024

Accepted: 24 February 2024

Published: 7 March 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: specialized lexicon; common lexicon; mental lexicon; wordnet; lexical semantics

1. Introduction

The specialized lexicon, i.e., words that denote more specific concepts and knowledge, emerging from specific domains such as chemistry, medicine, linguistics, etc., co-exist and co-occur with the common lexicon, i.e., the set of words that denote concepts and knowledge shared by average speakers, regardless of their specific training or expertise. The specialized lexicon develops from the need for new forms to convey new meaning, the latter being entirely new or only more precise. In this sense, the specialized lexicon aims at guaranteeing conciseness and semantic uniqueness (i.e., one form corresponding to one meaning) and at reducing ambiguity (Gotti 2003). The goal of domain-specific codes is to allow for a more informative and transparent communication between people sharing similar knowledge backgrounds, in other words, specialists. However, specialized lexicon is not used in the void and specialists do not become unable to use common language.

The level of interdisciplinarity in current fields of knowledge and daily activities, the growing interest in specialized and technical knowledge and the increasing access to information by non-specialists are changing our perspective on communication contexts, needs and linguistic models. This is the case for specialist/non-specialist communication, for instance, between health professionals and patients (Elhadad and Sutaria 2007; Smith and Fellbaum 2004); interdisciplinary work environments (e.g., artificial intelligence) with

heterogeneous teams and projects covering and combining different knowledge domains and subfields (Motos 2011); and diverse teaching–learning processes (Fuentes 2001) entailing different levels of expertise and explicit knowledge transfer needs. In these contexts, we activate different uses of the language and different lexica, depending on the target public and our communicational and informational needs (Cabr e 1999; Loukachevitch and Dobrov 2004).

Well-established literature (e.g., Pearson 1998; Cabr e 1999; Le on Ara uz et al. 2012) describes specialized languages as sub-codes of the general or common language, being general or common language the set of rules and the set of lexical units shared by a given linguistic community and used in “unmarked” situations. Specialized languages, on the other hand, are sub-codes that partially overlap with the general code but show some specificities, namely “subject field, type of interlocutors, situation, speaker’s intentions, the context in which a communicative exchange occurs, the type of exchange, etc.” (Cabr e 1999, p. 59), and, we add, specific semantic features that mirror specialized conceptualizations. Le on Ara uz et al. (2012) confirm this approach, stating that “differences mainly involve the predominance of a set of conceptual categories and relations”. How these co-exist in one single (mental or digital) database is highly relevant for several research fields.

Current studies on psychology, psycholinguistics and learning analysis show that the mental lexicon (or lexical memory) functions as a network that, although individual to each speaker, for instance according to different degrees of expertise in different domains (Yee et al. 2018), displays a general topography common to all speakers (Huth et al. 2016). These studies corroborate early lexical semantics research on relational models of the lexicon (Miller et al. 1990), showing that knowledge and language are best modeled as language networks of words and concepts. Language models “provide a fruitful way of representing the complexity of the mental lexicon, (. . .). Using network science approaches to model knowledge networks respects the inherent complexity and vagueness of knowledge representations.” (Siew 2022, p. 121).

When applied to domain-specific knowledge, research on the network structure has shown that general and domain-specific semantic networks are connected and that the number of inter-relations between these networks increases their efficiency and the speakers’ semantic fluency, similarly to bilingual speakers (Siew and Guru 2023; Llach 2023). These findings explain specialists’ ability to “de-specialize” (popularize/vulgarize) their speech and their ability to use, access and switch to both general and specialized codes, as multilingual speakers can switch among different natural languages.

Thus, the issue of how these sub-codes are connected and communicate with each other is not only theoretically interesting and relevant for linguistics, language and communication, psychology and knowledge representation but also crucial for the development of resources for natural language processing (NLP) solutions concerning specialized and non-specialized communication.

Existing lexical semantic approaches, and particularly lexical semantic theories describing the lexicon and the relations established between lexical units, typically do not cover differences in sub-codes (general vs. specialized). Definitions of synonym or hypernym relations, for instance, despite considering context (e.g., in Fellbaum 1998 or in Vossen 2002), do not account for concept specialization or for the sub-coding linking/mapping that allows, blocks or mediates communication between specialist and non-specialist speakers. Yet we are able to verify that, on many occasions, specialists and non-specialists use different lexical units to talk about the same thing, successfully conveying meaning and exchanging information.

Using the WordNet model (Miller et al. 1990; Fellbaum 1998; Vossen 2002) as our base framework, in this paper, we analyze the relations between specialized and common lexical units, focusing on the reference potential lexical units can be associated with. Our proposal results in a regular model for linking common and specialized lexica, accounting for specialized and non-specialized lexica in a single database, described and interconnected in a formal way. The resulting resources can be used in several NLP tasks, such as information

extraction, semantic disambiguation, document indexing and retrieval, etc., but also for improving specialist/non-specialist communication. Integrated and comprehensive resources can be very useful for dealing with communicative contexts requiring or addressing mixed codes, such as teaching/learning situations and health professional–patient interactions, where code-switching is frequent and necessary.

2. Methods and Framework

The WordNet model (Miller et al. 1990; Fellbaum 1998) is a formal lexical semantics model that reflects the organization of the mental lexicon. Wordnets (i.e., resulting lexical resources built according to the model) are computational relational databases based on the notion of concept, the core element of the network (Fellbaum 1998). Concepts can be lexicalized by a single lexical unit or by several synonyms, grouped in synsets (i.e., synonym sets), and the meaning of each set is represented and described through the relations it establishes with others in the network, including through its position in the hierarchy. This approach, in which lexical knowledge is seen as an intermediate level between language and concepts, follows neo structuralist semantics (Geeraerts 2010).

Wordnets establish a lexical hierarchy through the hypernymy/hyponymy relation that is quite relevant for modeling information inheritance. The position of a node reflects its level in the inheritance chain, from more generic concepts (hypernym synsets) in the upper levels to more specific concepts (hyponym synsets) in the lower levels (Miller et al. 1990). Relational databases do not describe lexical units or the concepts they lexicalize as individual entities but function in a structural manner as pieces of a larger puzzle, as shown in Figure 1.

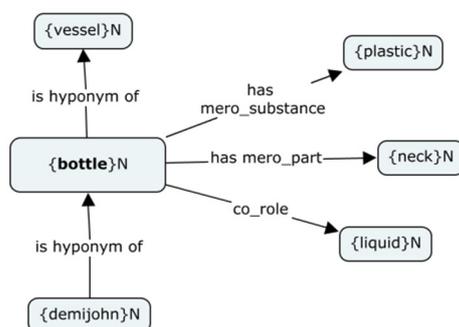


Figure 1. Wordnet nodes; legend: {} = synset; N = noun, V = verb, ADJ = adjective.

The WordNet model is based on three major notions:

- Word/lexical unit: linguistic unit that consists of a set of sounds/characters with morphological and syntactic properties with which a given stable meaning is associated.
- Meaning: linguistic unit that consists of set of semantic properties/values stably associated to a given word, reflecting prototypical conceptual knowledge, typically including a hypernym and specific and differentiating characteristics.
- Concept: knowledge unit that consists of a (mental) representation of a specific part of information, individualized by cognitive processes. Conceptual aspects concern, thus, the cognitive aspects relevant to the establishment and identification of a given portion of knowledge, and the conceptualization process concerns the cognitive process that defines the cognitive aspects and properties relevant to the delineation of a given concept.

Contrary to traditional lexical resources such as dictionaries, in the WordNet model, the unit (i.e., the node in the network) is the concept, represented by the set of words that can lexicalize it (synset). This is what makes it a lexical conceptual model. The synonymy relation used in the WordNet model corresponds to a relation among words that convey semantic similarity or synonymy in context, defined by Miller et al. (1990, p. 6) as “two expressions are synonymous in a linguistic context C if the substitution of one for the other

in C does not alter the truth value". It is a symmetric relation in which if A and B are synonyms in C, then A is B in C and B is A in C. Synonymy is, thus, a structural semantic relation of the model, establishing its foundation. In further versions of the model, such as the EuroWordNet model, other conditions apply to synonymy relation. Namely, two words are synonyms if

(i) They "denote the same range of the entities, irrespective of the morpho-syntactic differences, differences in register, style or dialect or differences in pragmatic use of the words, (. . .).

(ii) they cannot be related by any of the other semantic relations defined" (Vossen 2002, p. 18).

As it occurs in the mental lexicon, words can be part of as many units (lexical conceptual nodes) as the concepts they can lexicalize: *bank*₁—type of financial institution; *bank*₂—slope of land; *bank*₃—supply or stock; *bank*₄—bench for rowers; . . .

The other structuring relation, the hypernym/hyponymy relation, can be defined as follows:

1. X is **hyponym** of Y if X is a type of Y and Y is not a type of X.

Hyponymy/hypernymy is a lexical conceptual relation that concerns both world knowledge and linguistic knowledge. This can be tested by lexical anaphoric constructions where the hypernym is used to refer a more specific referent (the hyponym) previously introduced.

2. a. He bought a pitbull, but the dog doesn't bite.
- b. He crawled through the woods, moving so to avoid being seen by the guards.
- c. Marine animals can be endangered by the near proximity of cities since many aquatic animals are easily affected by sewage pollution (adapted from Amaro 2009, p. 26).

Lexical units are organized according to their type, a hyponym being all that its hypernym is and more. This relation contemplates the definition of a monotonic inheritance device (see Miller et al. 1990) that allows adequate and economic descriptions, since hyponyms inherit the properties of their hypernym.

So, synsets (the sets of synonym lexical units that can be used to refer to a specific concept) are related to each other through diverse types of relations: lexical conceptual relations (hypernymy/hyponymy, meronymy, etc.), function or role relations (agent-patient relation, involved instrument relation, etc.), semantic opposition relations (antonymy, near antonymy) and cause relations (is caused by/causes, etc.), all formally defined and linguistically validated (see Amaro et al. 2013).

The formal aspect of this framework assures that subsets of any lexicon, including specialized ones, are described in the same relational manner through labeled lexical semantic relations established between synsets that allow for describing meaning, inheritance and inference properties. Thus, the merging and integration of specialized and non-specialized lexica is expected to be possible and adequately modeled.

To define the proposal depicted in this paper, we explored existing related work on the integration of domain-specific lexical networks in general wordnets and examined the notion of reference in a semantic pragmatic approach to determine if a proximity relation can be formally defined and implemented in the model to connect units from common and specialized domains. This includes the definition of relations for nominal nodes, verbal nodes and adjectival nodes, considering the major word classes that convey meaning in languages (including specialized domains). The relations are formally defined in terms of symmetry and implementation restrictions, using the available model apparatus. Validation tests are also foreseen.

3. A New Relation for Integrating Common and Specialized Lexica

Departing from the notion of the reference potential of expressions, further explored in the discussion section (Section 4), we propose a relation that links lexical conceptual nodes (synsets) from common and specialized lexica, modeling their semantic similarity.

Common lexica and specialized nodes cannot be defined as synonyms, since they denote different intentional properties (therefore, different or partially different meanings), which in wordnets are represented by sets of different relations established with other nodes in the network. However, analogously to units from different natural languages, they can point to a same potential, generic and prototypical referent. This is what makes these synsets equivalents, in the sense that they can be used to talk about the same thing, even though the concepts they denote, and thus the semantic properties that make up their meaning, necessarily differ, explaining the need for the creation of specialized units. Therefore, understanding, making explicit and modeling this degree of equivalence between different codes contributes to avoid miscommunication phenomena between experts and non-experts.

We adopt and adapt the notion of generic reference described by [Cruse \(2000\)](#), which concerns the reference to a class of entities interrelated to the semantic potential value of that expression to refer to something, and we add it to the set of notions presented above:

- Reference potential: the ability of words to indicate specific parts/elements of the real world. Therefore, the referent is the part/element of the real world that a given word indicates, when used in a defined context.

This pragmatic aspect accounts for the use of the expressions in an enunciative context, while the semantic aspect ensures the independence and systematicity of the representation. This is a relevant theoretical distinction since, in a lexical conceptual network, we represent meaning or point to prototypical entities independently of any enunciative context.

3.1. Inter-Code Equivalence

Previous studies distinguish between three kinds of lexical conceptual variations in WordNet model that result in three different strategies for merging synsets from specialized and common networks ([Amaro and Mendes 2012](#)):

- Compatible synsets: the specialized synset denotes a concept similar to the one denoted by the common node, but more precise. The hypernym chain is the same for both synsets, but the specialized synset establishes more horizontal relations.
- Semi-compatible synsets: the specialized hypernym chain has an intermediary hypernym expressing some specification about the specialized concept that does not exist in the common lexicon network.
- Incompatible synsets: the specialized and common nodes, though related, do not share elements in the hypernym chain, directly or indirectly. In that case, there is no possible merging.

The case presented in Figure 2 as incompatible is modeled according to [Amaro and Mendes \(2012, p. 152\)](#)¹. Thus, synsets $\{attic_1, loft, garret\}_N$ (\approx low floor immediately under the roof), and $\{attic_2\}_N$ (\approx low story immediately under the entablature or cornice), show

- Two intentional denotations partially overlapping (i.e., the two different sets of features that determine the meaning of the lexical unit, corresponding to the sets of lexical conceptual relations each node establishes with the other nodes in the network);
- Two different sets of synonyms that can refer to the class of entities that satisfy the sense. In other words, two ways of mentioning the concept at stake, respectively: $\{attic, loft, garret\}$ and $\{attic\}$.

According to [Amaro and Mendes \(2012\)](#), in an integrated resource, these two nodes would be independent one from the other, as in, for instance, a typical case of regular polysemy with incompatible meanings. However, the nodes share several salient features: the same indirect holonym, $\{building\}_N$; the same characteristic, $\{low\}_{ADJ}$; and close role relations ($\{entablature\}_N$ is related to $\{roof\}_N$).

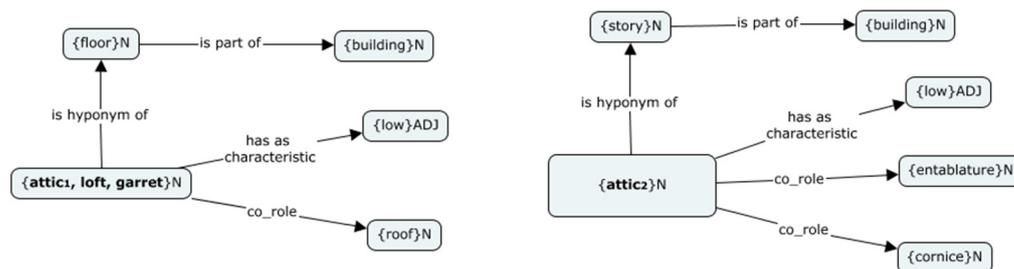


Figure 2. $\{attic_1, loft, garret\}_N$ common noun synset and $\{attic_2\}_N$ specialized noun synset from the architecture domain.

As these are in fact not completely unrelated and can point to the same entity in the real world, we hypothesize that this semantic proximity can be represented through a new relation, defined in informal terms as follows:

3. **INTER-CODE EQUIVALENCE:** Relation established between two synsets from common and specialized lexica with the same part of speech (POS), that stands for a relevant degree of sharing of semantic and conceptual properties, reflected by the sharing of the hypernymy chain (directly or indirectly) and of other horizontal relations.
4. **X IS INTER-CODE EQUIVALENT TO Y** if
 - a. X is a synset from common lexicon and Y is a synset from a specialized lexicon;
 - b. X is not hypernym/hyponym of Y;
 - c. X and Y establish, directly or indirectly, hierarchical and horizontal relations with the same synsets or with synsets that are inter-code equivalents;
 - d. X and Y can be replaced one for the other in a given context C, without changing the truth value of C.

Example: $\{attic_1, loft, garret\}_N$ **IS INTER-CODE EQUIVALENT TO** $\{attic_2\}_N$

Stating these conditions for the existence of the inter-code equivalence relations, we can test them in real examples, as shown below.

5.
 - a. $\{attic_1, loft, garret\}_N$ belongs to common lexicon and $\{attic_2\}_N$ belongs to specialized lexicon—TRUE;
 - b. $\{attic_1, loft, garret\}_N$ IS HYPONYM/HYPERONYM OF $\{attic_2\}_N$ —FALSE;
 - c.(i) $\{attic_1, loft, garret\}_N$ IS PART OF $\{building\}_N$ (indirectly) and $\{attic_2\}_N$ IS PART OF $\{building\}_N$ (indirectly) as well—TRUE;
 - c.(ii) $\{attic_1, loft, garret\}_N$ HAS AS CHARACTERISTIC $\{low\}_{ADJ}$ and $\{attic\}_N$ HAS AS CHARACTERISTIC $\{low\}_{ADJ}$ as well—TRUE;
 - c.(iii) $\{attic_1, loft, garret\}_N$ CO-ROLE $\{roof\}_N$ and $\{attic_2\}_N$ CO-ROLE $\{roof\}_N$ (indirectly) as well—TRUE;
 - d. If ‘The renovation damaged the attic/loft/garret!’ is true, then ‘The renovation damaged the attic!’ is true—TRUE.

The sharing of some of the meaning components between the nodes and the fact that it is possible to use the lexical units in them to refer to the same entity in the real world attest to the semantic similarity that relates the nodes, similarly to what happens in synonymy and equivalence in different languages.

In the case shown in (5), the relation between the common and specialized synsets is quite obvious, and it could also be explained by polysemy if we shift from an onomasiological to a semasiological approach. However, even from a semasiological perspective, the issue of how the specialized and non-specialized nodes are connected would remain. The case above still fulfills the conditions required, and it is possible to represent this similarity through an inter-code equivalence relation between the two nodes that maintains all meaning characteristics.

Figure 3 presents a distinct case in which speakers may be able to recognize the process of semantic extension that originated the new meaning of *virus* in the computer science

domain, i.e., the fact that it is harmful to the host and is able to and uses the hosts' abilities to replicate.

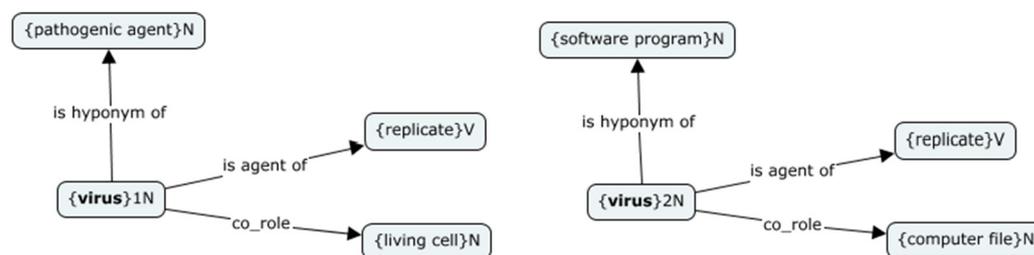


Figure 3. *{virus}1_N* common noun synset and *{virus}2_N* specialized noun synset from the Computer Science domain.

However, there is no overlapping in this case in terms of (i) meaning—there is no contact in the hypernymy chains of the two synsets, nor is there other shared relations—nor in terms of (ii) contexts of use (domain, co-occurrences, etc.) in which a potential referent (i.e., the real-world entity we are talking about) could be the same. This consists, thus, in a case of irregular polysemy or homonymy.

Following the methodology presented in (5), we can apply the same test to *{virus}1_N* and *{virus}2_N*, demonstrating that there is no overlap in this semantic extension process:

6. a. *{virus}1_N* belongs to common lexicon and *{virus}2_N* belongs to specialized lexicon;
- b. *{virus}1_N* IS HYPONYM/HYPERONYM OF *{virus}2_N*—FALSE;
- c.(i) *{virus}1_N* IS AGENT OF *{replicate}V* and *{virus}2_N* IS AGENT OF *{replicate}V* as well—TRUE;
- c.(ii) *{virus}1_N* CO-ROLE *{living cell}N* and *{virus}2_N* CO-ROLE *{living cell}N* as well—FALSE
- c.(iii) *{virus}2_N* CO-ROLE *{computer file}N* and *{virus}1_N* CO-ROLE *{computer file}N* as well—FALSE
- d. If ‘Contracting the virus₁ in the last century could be mortal and lead to the extinction of the species.’ is true, then ‘Contracting the virus₂ in the last century could be mortal and lead to the extinction of the species’ is also true—FALSE.

The condition in (6)d. can be further tested by resorting to lexical anaphora to assure that the proper lexical unit is being used. For instance: if ‘Contracting the virus₁ in the last century could be mortal and lead to the extinction of the species; nowadays it (the virus₁) is not because we found a way to destroy the pathogenic agent.’ is true, then ‘Contracting the virus₂ in the last century could be mortal and lead to the extinction of the species; nowadays it (the virus₂) is not because we found a way to destroy the software program.’ would also need to be true, but is false.

Unlike the example in (5), the case above fulfills just one of the four conditions required. This means that the inter-code equivalence relation cannot link these two nodes.

The same relation can be established between the nodes *{weight}N* (≈physical property of an object concerning its heaviness) from common lexicon and *{mass}N* (≈physical property of a body concerning its matter) from specialized lexicon, as shown in Figure 4. In this case, the truth conditions are fulfilled even if the synsets do not establish horizontal relations with the same nodes directly but with nodes that are inter-code equivalents (*{matter}N*—*{heaviness}N*; *{body}N*—*{object, material entity}N*), and the replacement in the same contexts is possible: if ‘He brought a rock with the weight of 3 kg.’ is true, than ‘He brought a rock with the mass of 3 kg.’ is also true is TRUE.

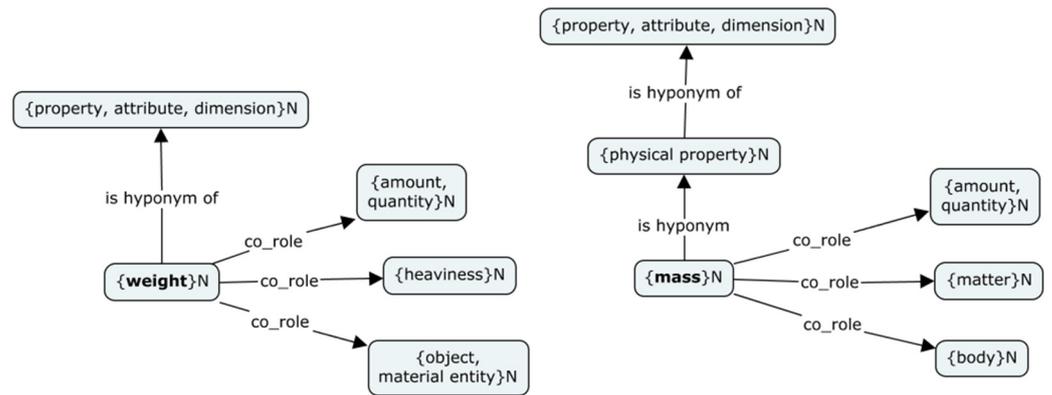


Figure 4. $\{weight\}_N$ common noun synset and $\{mass\}_N$ specialized noun synset from the physics domain.

If we consider further a comprehensive lexicon integrating several (or potentially all) specialized domains, the need to link nodes may cross different domains simultaneously. The next example in Figure 5 shows two different definitions from two different specialized domains.

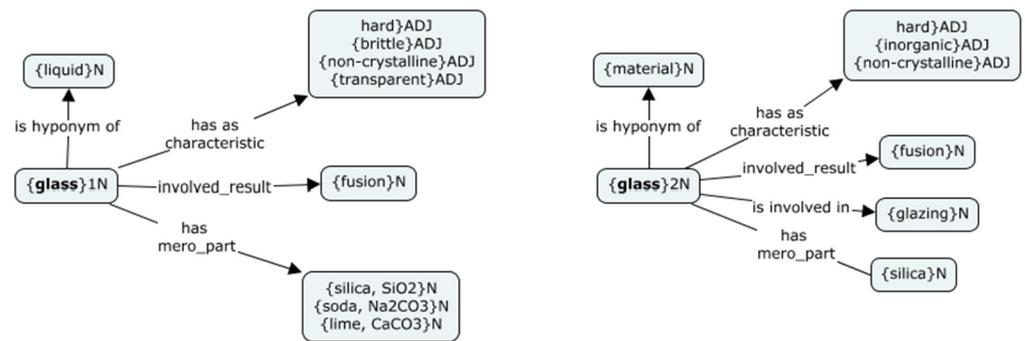


Figure 5. $\{glass\}_{1N}$ specialized noun synset from the chemistry domain and $\{glass\}_{2N}$ specialized noun synset from the materials engineering domain.

Following the tests determined above for common and specialized nodes, we applied the same method to two specialized nodes belonging to different domains. The results are shown in (7).

7. a. $\{glass\}_{1N}$ belongs to the specialized domain of chemistry and $\{glass\}_{2N}$ belongs to the specialized domain of materials engineering—TRUE;
- b. $\{glass\}_{1N}$ IS HYPERNYM/HYPONYM OF $\{glass\}_{2N}$ —FALSE;
- c.(i) Both $\{glass\}_{1N}$ and $\{glass\}_{2N}$ HAS AS CHARACTERISTIC $\{non-crystalline\}_{ADJ}$ and $\{transparent\}_{ADJ}$ —TRUE;
- c.(ii) Both $\{glass\}_{1N}$ and $\{glass\}_{2N}$ INVOLVED_RESULT $\{fusion\}_N$ —TRUE;
- c.(iii) Both $\{glass\}_{1N}$ and $\{glass\}_{2N}$ HAS_MERO_PART $\{silica\}_N$ —TRUE;
- d. If ‘Modern buildings use $glass_1$ to take advantage of natural light (since this liquid is transparent).’ is true, then ‘Modern buildings use $glass_2$ to take advantage of natural light (since this material is transparent).’ is also true—TRUE.

The example above shows that, even if each domain has a different focus according to the nature of the domain under analysis (León Araúz et al. 2012), the referent of both specialized expressions can be the same. Consequently, we can state that $\{glass\}_{1N}$ IS INTER-CODE EQUIVALENT TO $\{glass\}_{2N}$. Experts in communication point out that one of the main challenges within interdisciplinary teams, both in academic or non-academic contexts, is the use of specialized lexicon or jargon (Guo et al. 2023; Winowiecki et al. 2011). It seems to us that, besides the need for specialized lexicon resources and tools,

there is also the need for platforms combining different lexica and highlighting the degree of meaning equivalence between units from diverse knowledge domains. In this sense, establishing bridges among different codes, through semantic/pragmatic relations, will improve successful communication among heterogeneous interlocutors.

The INTER-CODE EQUIVALENCE relation proposed here accounts, thus, for the integration of specialized domains nodes in common lexicon, even considering different domains simultaneously. The next subsection presents the application of this relation to synsets from other relevant parts of speech, without any loss of adequacy.

Inter-Code Equivalence in Verbal and Adjectival Nodes

Although not as frequent as nouns in specialized domains, verbs and adjectives can also express domain-specific concepts (L'Homme 2002; L'Homme 2007; López Rodríguez 2007). As corroborated by recent studies on mental lexicon networks, nouns are the lexical category that shows a clearer hierarchical and taxonomic structure, with (Qiu et al. 2021), while verbs and adjectives are organized in more flat, less condensed and less modular networks (Qiu et al. 2021; Miller et al. 1990; Fellbaum 1998). However, the WordNet model allows for rich and adequate representations of these categories through the use of other relations available besides the hypernymy/hyponymy relation.

This section shows how the INTER-CODE EQUIVALENCE relation proposed can apply to verbal and adjectival nodes, adequately modeling the semantic similarity between different conceptualizations lexicalized by these categories in general and specialized domains.

- Verbs.

The verbal lexicon can be classified and organized according to lexical aspectual and internal properties that roughly correspond to Vendler's classic Aktionsart properties (Vendler 1967). Considering these properties, verbs can be described as denoting (i) states, (ii) processes or (iii) activities and transitions (Amaro 2009; Pustejovsky 1995). In a very simplified way, (i) states, when denoted by adjectives, contemplate static situations, such as *believe* or *exist*; (ii) processes typically display regular and non-limited dynamic situations, such as *write* or *swim*; and (iii) transitions are complex situations in which a process leads to or causes a final state different from an initial one such as *die* or *sink*. The fact that all these situations involve typical participants and conditions of realization (i.e., arguments) also allows us to better conceptualize them (Baker 2014; Vossen et al. 2018).

Considering these conceptual descriptions, verbs, as their correspondent deverbal nouns, i.e., nouns that derive from a verbal stem and that denote the same event, can be used to talk about the same situations: "The POS difference leads to subtle differences in meaning (such as argument reduction of nominalizations), but in many cases languages offer a choice between a noun, verb or adjective to name the same situation or event." (Vossen 2002, p. 19). This means that, as nouns, other POSs have the ability to refer to entities and situations of the external world such as, for instance, $\{move\}_V$ and $\{movement\}_N$, $\{paint\}_V$ and $\{painting\}_N$, $\{destroy\}_V$ and $\{destruction\}_N$, etc., and can thus be analyzed as nouns in terms of the equivalence relation established. The analyses carried out by Correia (2002) make this explicit and underline the consequent parallelism in the syntactic realization between nominalizations (deverbal nouns) and verbs. Following Anscombe (1986), the author presents a noun typification ranging from processive/progressive nouns to stative nouns, action nouns, activity nouns, cyclic nouns, resultative nouns and cyclic resultative nouns. The proposal states that both the verb and its correspondent nominalization can support the same sentence structure, as shown in the examples below:

8. a. 'X demonstrates the problem for two hours'—'The demonstration of the problem occurs 8 for two hours' → A cyclic resultative verb supports an adverbial expression of duration—so does the deverbal nominalization.
- b. * 'X solves the problem for two hours'—* 'The solution of the problem occurs for two 11 h' → A resultative verb does not support an adverbial expression of duration, and neither does the deverbal nominalization.

We can examine, for instance, the case of the activity verb *breathe*, where we are able to identify a potential referent for the situation denoted by the deverbal noun *breathing*; we must necessarily be able to do the same for the correspondent verb, since the situation and the participants of the event are exactly the same, as shown in Figure 6.

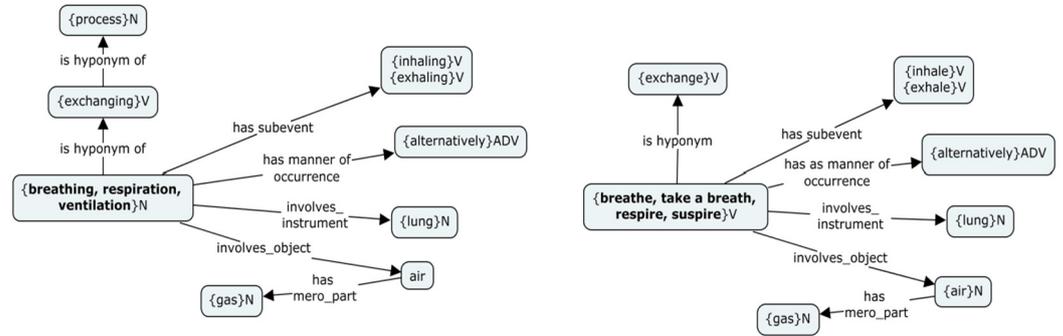


Figure 6. Deverbal noun representation vs. verb representation.

We can, then, proceed to test the applicability of the INTER-CODE EQUIVALENCE relation to verbal nodes, as we did for nouns. Let us take the examples $\{breathe_1, take\ a\ breath, respire_1, suspire\}_V$ (\approx exchange air alternatively by inhaling and exhaling using lungs) and $\{breathe_2, respire_2\}_V$ (\approx exchange gases by alternatively consuming oxygen and producing carbon dioxide through lungs or gills) from the common lexicon and specialized lexicon of Biology domains, respectively.

Here, it is apparent that the (minimum and sufficient) overlapping conditions between the common and the specialized nodes are satisfied:

- Directly, e.g., HYPERONYM relation and INVOLVES_INSTRUMENT relation;
- Indirectly, e.g., INVOLVES_OBJECT relations.

Thus, the example in Figure 7 shows that, even if the way of conceiving the breathing event is partially different since the specialized net is more detailed and precise, the situation evoked, and consequently the potential referent, can be the same. It is possible, then, to apply and test the INTER-CODE EQUIVALENCE relation to these nodes.

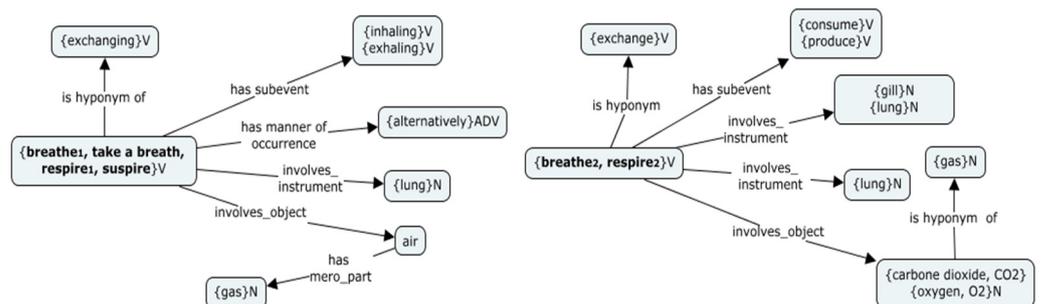


Figure 7. $\{breathe_1, take\ a\ breath, respire_1, suspire\}_V$ common verb synset vs. $\{breathe_2, respire_2\}_V$ specialized verb synset from the biology domain.

9. $\{breathe_1, take\ a\ breath, respire_1, suspire\}_V$ IS INTER-CODE EQUIVALENT TO $\{breathe_2, respire_2\}_V$ if
 - a. $\{breathe_1, take\ a\ breath, respire_1, suspire\}_V$ belongs to common lexicon and $\{breathe_2, respire_2\}_V$ belongs to the specialized lexicon of the biology domain;
 - b. $\{breathe_1, take\ a\ breath, respire_1, suspire\}_V$ IS HYPONYM/HYPERONYM OF $\{breathe_2, respire_2\}_V$ —FALSE;
 - c.(i) $\{breathe_1, take\ a\ breath, respire_1, suspire\}_V$ IS HYPONYM OF $\{exchange\}_V$ and $\{breathe_2, respire_2\}_V$ IS HYPONYM OF $\{exchange\}_V$ as well—TRUE;

- c.(ii) {*breathe*₁, *take a breath*, *respire*₁, *suspire*}_V INVOLVES INSTRUMENT {*lung*}_N and {*breathe*₂, *respire*₂}_V INVOLVES INSTRUMENT {*lung*}_N as well—TRUE;
 - c.(iii) {*breathe*₁, *take a breath*, *respire*₁, *suspire*}_V INVOLVES OBJECT {*air*}_N that HAS MERONYM {*gas*}_N and {*breathe*₂, *respire*₂}_V INVOLVES OBJECT {*carbon dioxide*, *CO*₂}_N and {*oxygen*, *O*₂}_N that IS HYPONYM of {*gas*}_N as well—TRUE;
 - d. If ‘The patient died because she could not breathe/take a breath/respire/suspire.’ is true, then ‘The patient died because she could not breathe/respire.’ is also true—TRUE.
- Adjectives.

Following the categorization of the adjectives described in Mendes (2009) into two main subclasses (descriptive and relational adjectives), it is possible to assume that, as a rule, adjectives ascribe a property or set of properties, status (more or less temporary) or features, corresponding roughly to states, to an entity, typically lexicalized by a noun.

It is therefore possible to assume that we can identify the potential, generic and prototypical referent that state adjectives point to. This can be conceptualized indirectly through the noun that expresses the attribute or property or a set of properties expressed by the adjective at stake. For instance, for *green*, a color descriptive adjective, we are able to identify the state it denotes and can refer to the nominal attribute it evokes, the color green.

The example presented in Figure 8 shows that “the relation between adjectives and a given attribute is encoded in WordNet model resources by linking some adjectives—cluster focal adjectives—to nouns lexicalizing the relevant attribute, using the attribute relation.” (Mendes 2009, p. 92). In WordNet.PT, this relation is codified through the label “CHARACTERIZES WITH REGARD TO”, between {*green*}_{ADJ} and {*color*}_N, as shown in Figure 8.

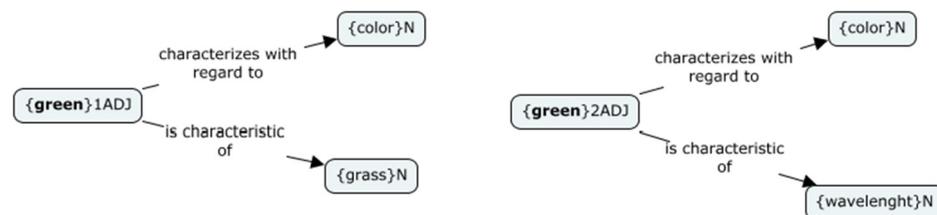


Figure 8. {*green*}_{1ADJ} common adjective synset vs. {*green*}_{2ADJ} specialized adjective synset from the physics domain.

Although the ways of conceiving the same adjective in the different sub-codes (common vs. specialized) are partially different, the referent activated for both synsets may be the same if both synsets share some relevant horizontal relations and if they can be replaced one for the other without changing the truth value of the sentence. That is, if we are talking about the same state. Consequently, when merging the two sub-codes, we can assert that: {*green*}_{1ADJ} IS INTER-CODE EQUIVALENT TO {*green*}_{2ADJ}.

The same is more straightforward for relational adjectives, as *maritime*, *geological*, *medical* etc., as “(.) relational adjectives are associated to a set of properties that, in general, corresponds to the denotation of a noun.” (Mendes 2009, p. 107). Figure 9 shows the inter-code equivalence relation between relational adjectives.

Another example illustrating the application of the INTER-CODE EQUIVALENCE relation in adjectival nodes is the adjective *melodic*, as described in WordNet 3.1². Considering the common domain and the specialized domain of music, the nodes in which *melodic* occurs are {*melodious*, *melodic*₁, *musical*}_{ADJ} (≈containing or constituting or characterized by pleasing melody) and {*melodic*₂}_{ADJ} (≈of or related to melody or tonal pattern (≈the perception of pleasant arrangements of musical notes)). These are represented in the WordNet model, as shown in Figure 10.



Figure 9. $\{\text{solar}\}_1\text{ADJ}$ common relational adjective synset vs. $\{\text{solar}\}_2\text{ADJ}$ specialized relational adjective synset from the astronomy domain.

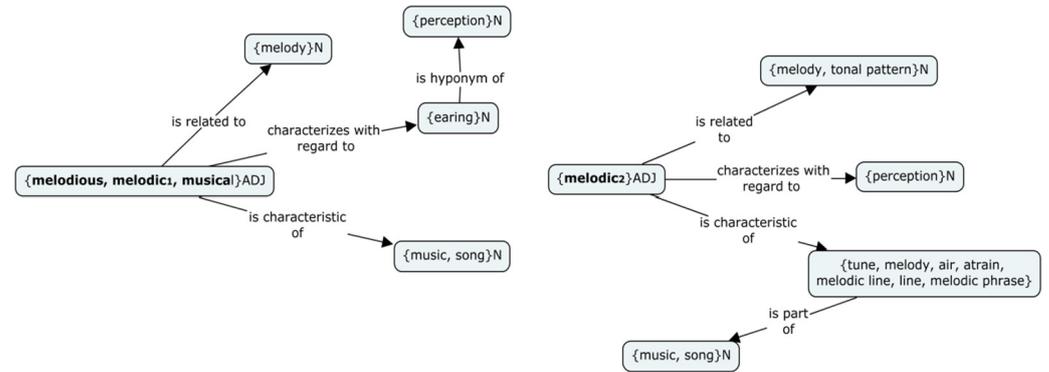


Figure 10. $\{\text{melodious, melodic}_1, \text{musical}\}_\text{ADJ}$ common adjective synset vs. $\{\text{melodic}_2\}_\text{ADJ}$ specialized adjective synset from the music domain.

Finally, as previously stated, adjectives denote states, consensually described as a kind of abstract entity, a type of situation, as represented in Figure 11.

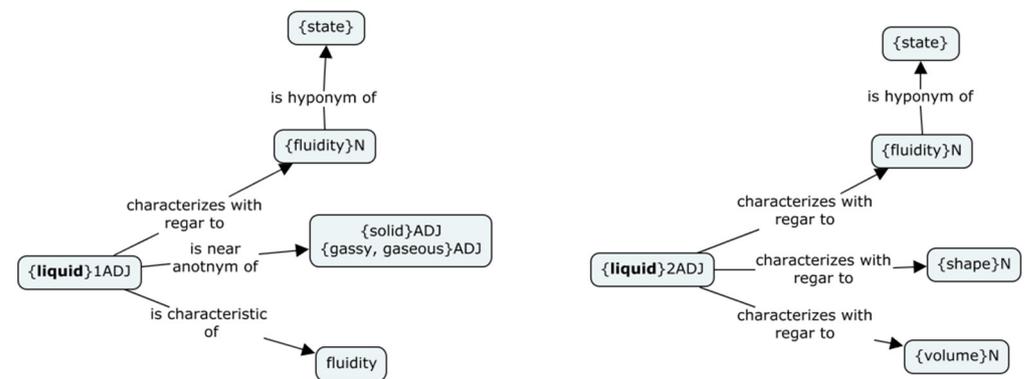


Figure 11. $\{\text{liquid}\}_1\text{ADJ}$ common adjective synset vs. $\{\text{liquid}\}_2\text{ADJ}$ specialized adjective synset from the chemistry domain.

As shown in the examples presented, adjective nodes are not as densely related to other nodes as, for instance, nominal nodes, resulting in less dense networks of relations. As it happens with verbs, adjectival networks are less hierarchical. In most of the cases, it is not possible to establish hypernym or hyponym relations, and the meaning of a given adjective node is expressed through other relations. However, the examples above sustain that our proposal to apply the INTER-CODE EQUIVALENCE relation to adjectives is viable and testable.

10. $\{\text{melodious, melodic}_1, \text{musical}\}_\text{ADJ}$ IS INTER-CODE EQUIVALENT TO $\{\text{melodic}_2\}_\text{ADJ}$ if
 - a. $\{\text{melodious, melodic}_1, \text{musical}\}_\text{ADJ}$ belongs to common lexicon and $\{\text{melodic}_2\}_\text{ADJ}$ belongs to the specialized lexicon of the music domain;
 - b. $\{\text{melodious, melodic}_1, \text{musical}\}_\text{ADJ}$ IS HYPONYM/HYPERONYM OF $\{\text{melodic}_2\}_\text{ADJ}$ —FALSE;

c.(i) {*melodious, melodic₁, musical*}_{ADJ} IS RELATED TO {*melody*}_N and {*melodic₂*}_{ADJ} IS RELATED TO {*melody, tonal pattern*}_N and {*melody*}_N IS INTER-CODE EQUIVALENT TO {*melody, tonal pattern*}_N—TRUE;

c.(ii) {*melodious, melodic₁, musical*}_{ADJ} CHARACTERIZES WITH REGARD TO {*earing*}_N that IS HYPONYM OF {*perception*}_N and {*melodic₂*}_{ADJ} CHARACTERIZES WITH REGARD TO {*perception*}_N—TRUE;

c.(iii) {*melodious, melodic₁, musical*}_{ADJ} IS CHARACTERIC OF {*music, song*}_N and {*melodic₂*}_{ADJ} IS CHARACTERIC OF {*tune, melody, air, strain, melodic line, line, melodic phrase*}_N that IS MERONYM OF {*music, song*}_N—TRUE;

d. If ‘These new songs are less melodious/melodic₁/musical.’ is true, then ‘These new songs are less melodic₂.’ is also true—TRUE.

As demonstrated in this subsection, the application of the INTER-CODE EQUIVALENCE relation to verbal and adjectival nodes is possible and motivated. This is an important contribution to guarantee the uniformity of relations mediating all the POSs encoded in relational lexical resources such as wordnets, on the one hand, and a simple and adequate way of accounting for all the main POSs occurring in specialized lexica—nouns, adjectives and verbs (L’Homme 2007; López Rodríguez 2007)—on the other.

4. Discussion

The proposal presented in the preceding section reflects the analysis of previous experiences integrating domain-specific networks in existent wordnets. In general, the idea that it is possible to establish some sort of equivalence relation between specialized and common nodes seems to be widely accepted (Freihat et al. 2013).

In the WordNet model, considering the general lexicon case, we can observe several situations³, namely

(i) One concept expressed by more than one word form/expression, which results in a single node composed by multiple lexical items related to each other by synonymy (corresponding to a synset), as shown in (11).

11. {*car, auto, automobile, machine, motorcar*}_N (≈motor vehicle with four wheels, usually propelled by an internal combustion engine).

(ii) Irregular polysemy or homonymy cases in which a same word form/expression expresses more than one concept, without any relation between them (corresponding to two or more synsets), as exemplified in (12).

12. a. {*bank₁*}_N (≈sloping land, especially the slope beside a body of water);

b. {*depository financial institution, bank₂, banking concern, banking company*}_N (≈financial institution that accepts deposits and channels the money into lending activities).

(iii) Regular polysemy with incompatible meanings (Pustejovsky 1995; Buitelaar 1998), resulting in more than one related node but not retrievable simultaneously or in the same contexts, as illustrated in (13)c.

13. a. {*spoon₁*}_N (≈piece of cutlery with a shallow bowl-shaped container and a handle);

b. {*spoon₂*}_N (≈amount that a spoon will hold);

c. #The silver butter spoon should be added to the boiling water.

(iv) Regular polysemy with compatible meanings, resulting in a single node with multiple hypernyms that aggregate the related meanings (e.g., building/institution), which in turn are retrievable simultaneously or in the same contexts, as illustrated in (14 b).

14. {*hospital*}_N (≈medical institution where sick or injured people are given medical or surgical care);

b. ‘The hospital that was robbed yesterday fired the security guards.’

A deeper analysis of the notions of synonymy and of regular polysemy is well presented in the literature (Buitelaar 1998; Copestake 1995; Copestake and Briscoe 1995; Cruse 1986) and is not the focus of this paper. However, these concepts are relevant to understand

how the notion of equivalence between specialized and common nodes is usually used and described in the related work described here and how relations are established between lexical conceptual representations.

As noticed in earlier work (Sagri et al. 2004; Chen et al. 2011; Amaro and Mendes 2012; Pedersen et al. 2012), sense discrimination covering specialized and common lexica often results in polysemy and semantic overlapping, and the way to achieve this requires new approaches to the available lexical semantic/lexical conceptual relations.

For instance, the method used to integrate domain-specific networks in the ItalWordNet, the wordnet for Italian (Roventini et al. 2000), based on the use of plug-in relations, allows for linking a specialized node to the common ‘equivalent’ through a kind of synonymy relation. This is the case of the plug-in synonymy relations for Eco-WN (ecology domain WordNet) and Jur-WN (law domain WordNet) (Magnini and Speranza 2001; Sagri et al. 2004), the integrative plug-in relations in Archi-WN (architecture domain wordnet) (Bentivogli et al. 2004) and the equivalent plug-in relations for the wordnet for maritime domain. The result is that domain-specific lexicalizations are integrated in common lexicon synsets.

By the principle of ‘specific-first,’ the resultant synset maintains the variants from the specialized synset, as well as the downwards and horizontal links (hyponyms, meronyms, etc.), since it is supposed to reflect a more complex and comprehensive conceptualization, while the upward links (hypernym) are only pulled from the common node, as shown in Figure 12 below.

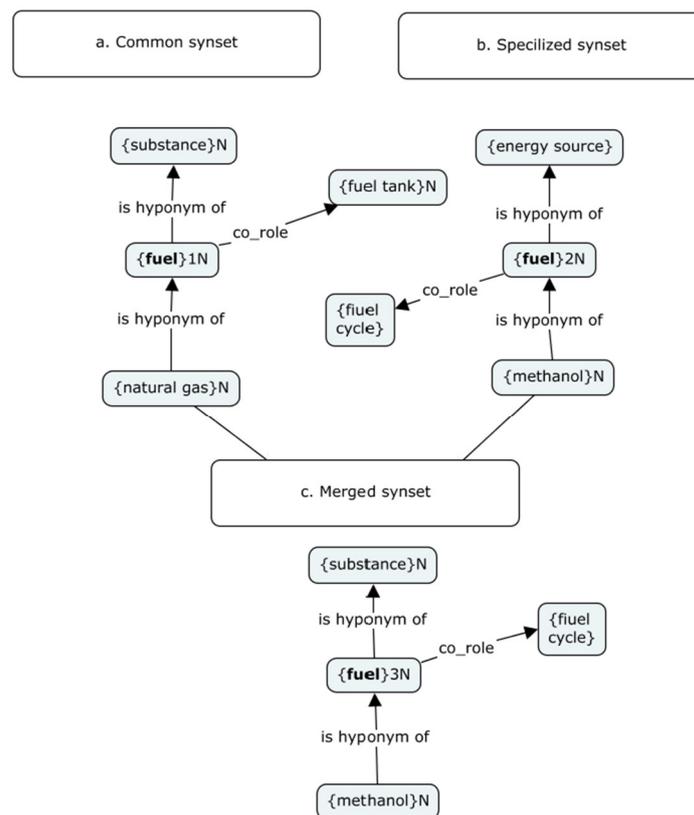


Figure 12. Common and specialized synset merging process.

Manual intervention in this process is reduced to the resolution of inconsistencies (Magnini and Speranza 2001; Roventini and Marinelli 2004), and so this method seems to be quite efficient and economic. However, the equivalence/synonymy notion used is theoretically problematic for several reasons.

First, if the nodes were really synonymous, they would not establish different lexical conceptual relations with other elements of the network, but they would simply overlap or

constitute larger synsets. Second, according to the organization and mapping of semantic information in the WordNet model, synonymy is a lexical relation between word forms (Miller et al. 1990, p. 6) that represents similarity of meaning existing exclusively between the elements of the same synset, which cannot “be related by any of the other semantic relations defined” (Vossen 2002, p. 18). Finally, as displayed in Figure 12, the ‘specific-first’ method does not prevent the loss of relevant information concerning hierarchical axes and horizontal relations, both from the common and the specialized nets.

This particular example suggests a dichotomy in the hyponymy chain between a type of sub-specification of the common node—characterizing the hyponym constitutive aspect (i.e., its constituents or parts, such as material, weight)—, in contrast with a role sub-specification for the specialized node—characterizing the telic/role aspect of the hyponym—, which is in fact a recurrent and common phenomenon concerning meaning “specialization” (e.g., *knife*—object for cutting), as explored in Pustejovsky (1995), Mendes and Chaves (2001) and Freihat et al. (2013).

Thus, in this case, we are not properly dealing with synonym synsets and therefore cannot establish a proper synonymy relation, as defined by the model, between the elements in the synset. Nonetheless, as we have demonstrated in our proposal, the synsets are linked through some kind of semantic similarity.

Following the proposal of Amaro and Mendes (2012), and, to a certain extent, Smith and Fellbaum (2004), our proposal looks at the notion of semantic equivalence, not strictly concerning the meaning of lexical units but also related to the referents that lexical units point to. In other words, two expressions coming from specialized and common lexica can have some differences in meaning and point to the same referent.

The example in Figure 13 distinguishes two ways of conceiving and representing the concept of ‘water’ through the different semantic properties encoded in the relations forming the network:

- different hypernyms: $\{liquid\}_N$ and $\{compound\}_N$;
- different horizontal relations, establishing different characteristics: $\{colorless\}_{ADJ}$; $\{odorless\}_{ADJ}$; $\{tasteless\}_{ADJ}$; meronyms $\{oxygen\}_N$; $\{hydrogen\}_N$; $\{covalent\ bond\}_N$; involvement relations expressing telicity $\{life\}_N$ and $\{solving\}_N$, according to the salience and relevance of the conceptual information for the speaker or for the domain.

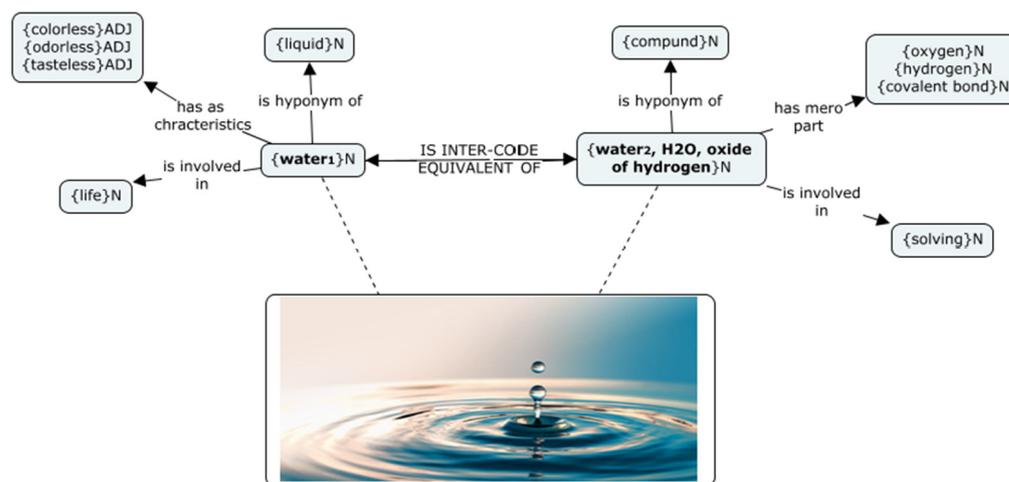


Figure 13. $\{water_1\}_N$ common synset and $\{water_2, H_2O, oxide\ of\ hydrogen\}_2N$ specialized synset from the chemistry domain, including the representation of their potential referent (image).

As León Araúz et al. (2012, p. 134) state, “[specialized lexical units] are lexical items, which can be understood and construed from different perspectives.” At the same time, however, the mental information/image/prototype that both nodes point to (Khoo and

Na 2006; Huth et al. 2016; Siew 2022; Siew and Guru 2023), can eventually be the same, as discussed in the previous sections.

In a different approach, and considering the topical context (i.e., contexts set according to the topic) as a link between co-referring expressions in different local contexts, ReferenceNet (Vossen et al. 2018) also posits that different synsets can, in specified contexts, point to the same entity. This idea is illustrated through a simple example: if we collect all the news articles about the same event, for instance a dispute or war, in different periodicals during the same week, we are expected to find different lexicalizations for the same participants of the event—*victim, aggressor, suspect, innocent, murder*, etc.—according to the existing points of view. As discussed in many semantic theories, meaning and reference are not necessarily coincident. Lexicalizations (words) may have different meanings, but they may refer to the same things or, in other words, talk about a same referent. In this way, different synsets belonging to the same hypernym chain are associated with the same ReferenceSet in ReferenceNet (Vossen et al. 2018). However, ReferenceNet only considers synsets from the same semantic domain, and it is not concerned with the connection or relation between specialized and non-specialized units, in particular when there is also semantic overlapping.

The efforts registered to merge specialized and non-specialized lexica through plug-in and proximity relations constitute strong motivation in the direction of straightforwardly modeling the connection between specialized and non-specialized nodes using the WordNet model to integrate different subsets of the lexicon. On the other hand, using the notion of reference together with strict conditions for semantic property sharing, as used in our proposal, provides a way to accurately express this relation. The relations between the nodes of wordnets (and the mental lexicon) and the relations established between words and entities in the real world are fundamental to understanding the meaning (and use) of the expressions themselves and, in particular, to describe and model a functional interface between specialized and common lexica.

5. Conclusions

In this paper, we propose a semantic pragmatic approach to linking common and specialized lexica in the WordNet model through a new INTER-CODE EQUIVALENCE relation based on the semantic properties and the reference potential of the lexical units involved. We define a solid and theoretically well-motivated relation that guarantees the uniformity of the relations encoded in wordnets while adequately and straightforwardly accounting for all the main POSs occurring in specialized lexica, maintaining the specialized meaning features. This proposal follows the basic observation that, even if specialized domains activate different sets of features or express concepts through different lexicalizations, there is often an overlap between specialized and common nodes and, in many cases, speakers use them concurrently, since the specialized and common nodes can point to the same reality.

Even though specialized networks are expected to be mostly populated by nominal nodes, we show that it is productive to test and extend the INTER-CODE EQUIVALENCE relation to the other POS (namely, adjectives and verbs) in a motivated and systematic way. Contrary to previous proposals, a regular model for bridging the gap between common and specialized lexica can thus be effectively established, accounting for specialized and non-specialized lexica in a single database and allowing for the entire lexicon of a language to be described and interconnected in a formal way.

The resulting resources can be used for tasks such as information extraction, disambiguation, document indexing and retrieval, etc., and also for improving and better understanding expert and non-expert communication, for instance by contributing to the ‘translation’ of specialized texts and information to non-specialists, objectively contributing to improving communication. Integrated or interrelated resources can also be very useful for dealing with communicative contexts requiring or addressing mixed codes at the same time, such as teaching–learning situations, health professional–patient interactions, efforts

towards the leveling of inequalities concerning access to information (e.g., financial literacy, legal and citizenship rights), interdisciplinary team interchanges where code-switching is frequent and necessary and where it is, in fact, essential to know if we are talking about the same things.

Moreover, the operationalization of the integration of sub-codes will allow the observation of relational patterns, which can be of great use for understanding the process of knowledge specialization and its reflection in the lexicon, organization, acquisition, maintenance, and growth (Wulff et al. 2019). Preliminary work on the data suggests possible patterns such as type–role alternation, providing some insights on the degree of overlap and dependency between nodes from different networks, as well as the level of domain dependency of the patterns. The study of these relations and relational patterns is also useful in avoiding biases as part of the lexicographic process of designing both common and specialized resources and definitions more unbiased.

Author Contributions: Conceptualization, methodology, formal analysis, investigation, C.B. and R.A.; writing—original draft preparation, data curation, visualization, C.B.; writing—review and editing, validation, supervision, R.A. All authors have read and agreed to the published version of the manuscript.

Funding: This research is supported by the Portuguese national funding through the FCT—Portuguese Foundation for Science and Technology, I.P. as part of the project UIDB/LIN/03213/2020; 10.54499/UIDB/03213/2020 and UIDP/LIN/03213/2020; 10.54499/UIDP/03213/2020—Linguistics Research Centre of NOVA University Lisbon (CLUNL) and by the PhD grant (PD/BD/128131/2016).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data from WordNet 3.1 can be accessed in <https://wordnet.princeton.edu> (accessed on 1 January 2023). Data from WordNet.PT are not yet available as they come from an ongoing PhD thesis.

Acknowledgments: We thank the anonymous reviewers for their constructive comments and suggestions.

Conflicts of Interest: The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

Notes

- ¹ Examples built based on the definitions and synsets presented in Amaro and Mendes (2012, p. 152), and in American Heritage® Dictionary of the English Language, Fifth Edition, 2016, and Collins English Dictionary—Complete and Unabridged, 12th Edition 2014 (<https://www.thefreedictionary.com/attic>), (accessed on 1 January 2023).
- ² Synsets and glosses retrieved from WordNet 3.1 (<https://wordnet.princeton.edu/>), (accessed on 1 January 2023).
- ³ Synsets in examples (11) to (14) were adapted from WordNet 3.1 (<https://wordnet.princeton.edu>). We uniformized the glosses by using the formula *direct hyperonym + specific differences*.

References

- Amaro, Raquel. 2009. Computation of Verbal Predicates in Portuguese: Relational Network, Lexical-Conceptual Structure and Context. Ph.D. thesis, University of Lisbon, Lisbon, Portugal.
- Amaro, Raquel, and Sara Mendes. 2012. Towards merging common and technical lexicon wordnets. In *Proceedings of the 3rd Workshop on Cognitive Aspects of the Lexicon*. Mumbai: The COLING 2012 Organizing Committee, pp. 147–60. Available online: <https://aclanthology.org/W12-5100> (accessed on 1 December 2023).
- Amaro, Raquel, Sara Mendes, and Palmira Marrafa. 2013. Increasing Density through New Relations and PoS Encoding in WordNet.PT. *International Journal of Computational Linguistics and Applications* 4: 11–27.
- Anscombe, Jean-Claude. 1986. Article zéro, termes de masse et représentation d'événements en français contemporain. *Recherches Linguistiques* 11: 5–34.
- Baker, Collin. 2014. FrameNet: A Knowledge Base for Natural Language Processing. In *Proceedings of 3 Frame Semantics in NLP: A Workshop in Honor of Chuck Fillmore (1929–2014)*. Baltimore: Association for Computational Linguistics, pp. 1–5. [CrossRef]

- Bentivogli, Luisa, Andrea Bocco, and Emanuele Pianta. 2004. ArchiWordNet: Integrating WordNet with Domain-Specific Knowledge. Paper presented at the 2nd International Global Wordnet Conference, Brno, Czech Republic, January 20–23; pp. 39–47. Available online: <https://www.fi.muni.cz/gwc2004/> (accessed on 1 December 2023).
- Buitelaar, Paul. 1998. CoreLex: An Ontology of Systematic Polysemous Classes. In *Proceedings of the 11 1st International Conference on Formal Ontology in Information Systems*. Trento: IOS Press, vol. 46, pp. 221–35. Available online: <http://www.coli.uni-saarland.de/publikationen/softcopies/Buitelaar:1998:COS.pdf> (accessed on 1 December 2023).
- Cabré, Maria Teresa. 1999. *Terminology: Theory, Methods and Applications*. Amsterdam and Philadelphia: John Benjamins Publishing Company.
- Chen, Rung Ching, Cho Tscan Bau, and Chun Ju Yeh. 2011. Merging domain ontologies based on the WordNet system and Fuzzy Formal Concept Analysis techniques. *Applied Soft Computing Journal* 11: 1908–23. [CrossRef]
- Copestake, Ann. 1995. Representing lexical polysemy. Paper presented at the AAI Spring Symposium on Representation and Acquisition of Lexical Knowledge Polysemy Ambiguity and Generativity, Palo Alto, CA, USA, March 27–29; pp. 21–22. Available online: <https://www.aai.org/Papers/Symposia/Spring/1995/SS-95-01/SS95-01-006.pdf> (accessed on 1 December 2023).
- Copestake, Ann, and Ted Briscoe. 1995. Semi-productive polysemy and sense extension. *Journal of Semantics* 12: 15–67. [CrossRef]
- Correia, Clara Nunes. 2002. Estudos de determinação. A operação de quantificação-qualificação em sintagmas nominais. In *Textos*. Lisbon: Fundação Calouste Gulbenkian e Fundação para a Ciência e a Tecnologia.
- Cruse, Alan. 1986. *Lexical Semantics*. Cambridge, UK: Cambridge University Press.
- Cruse, Alan. 2000. *Meaning in Language. An introduction to Semantics and Pragmatics*, 2nd ed. Oxford: Oxford University Press.
- Elhadad, Noemie, and Komal Sutaria. 2007. Mining a Lexicon of Technical Terms and Lay Equivalents. In *Proceedings of the Workshop on BioNLP 2007: Biological, Translational, and Clinical Language Processing*. Prague: Association for Computational Linguistics Stroudsburg, pp. 49–56. [CrossRef]
- Fellbaum, Christiane, ed. 1998. *Wordnet: An Electronical Database*. Cambridge, UK: MIT Press.
- Freihat, Abed Alhakim, Fausto Giunchiglia, and Biswanath Dutta. 2013. Solving Specialization Polysemy in WordNet. *International Journal of Computational Linguistics and Applications* 4: 29–52.
- Fuentes, Alejandro Curado. 2001. Lexical behaviour in academic and technical corpora: Implications for ESP development. *Language Learning & Technology* 5: 106–29.
- Geeraerts, Dirk. 2010. *Theories of Lexical Semantics*. Oxford: Oxford University Press.
- Gotti, Maurizio. 2003. *Investigating Specialized Discourse*. Lausanne: Peter Lang.
- Guo, Yue, Joseph Chee Chang, Maria Antoniak, Erin Bransom, Trevor Cohen, Lucy Lu Wang, and Tal August. 2023. Personalized Jargon Identification for Enhanced Interdisciplinary Communication. *arXiv* arXiv:2311.09481. [CrossRef]
- Huth, Alexander, G. Wendy A. de Heer, Thomas L. Griffiths, Frédéric E. Theunissen, and Jack L. Gallant. 2016. Natural speech reveals the semantic maps that tile human cerebral cortex. *Nature* 532: 453–58. [CrossRef] [PubMed]
- Khoo, Christopher S. G., and Jin-Cheon Na. 2006. Semantic Relations in Information Science. *Annual Review of Information Science and Technology* 40: 157–228. [CrossRef]
- León Araúz, Pilar, Pamela Faber, and Silvia Montero Martínez. 2012. Specialized language semantics. In *A Cognitive Linguistics View of Terminology and Specialized Language*. Edited by Pamela Faber. Berlin and Boston: De Gruyter Mouton, pp. 133–211.
- L’Homme, Marie-Claude. 2002. What can Verbs and Adjectives Tell us about Terms? Paper presented at the Proceedings of Terminology and Knowledge Engineering. 6th International Conference, Nancy, France, August 28–30.
- L’Homme, Marie-Claude. 2007. Using Explanatory and Combinatorial Lexicology to Describe Terms. In *Selected Lexical and Grammatical Issues in the Meaning-Text Theory. In Honour of Igor Mel’cuk*. Amsterdam and Philadelphia: John Benjamins Publishing Company, pp. 13–50.
- Llach, Maria Pilar Agustín. 2023. Mapping the mental lexicon of EFL learners: A network approach. *Revista de lingüística y lenguas aplicadas* 18: 1–17. [CrossRef]
- Loukachevitch, Natalia, and Boris Dobrov. 2004. Sociopolitical Domain As a Bridge from General 17 Words to Terms of Specific Domains. Paper presented at Second Global Wordnet Conference, Brno, Czech Republic, January 20–23; pp. 163–68. Available online: <https://www.fi.muni.cz/gwc2004/> (accessed on 1 December 2023).
- López Rodríguez, Clara Inés. 2007. Understanding scientific communication through the extraction of the conceptual and rhetorical information codified by verbs. *Terminology. International Journal of Theoretical and Applied Issues in Specialized Communication* 13: 61–84. [CrossRef]
- Magnini, Bernardo, and Manuela Speranza. 2001. Integrating Generic and Specialized Wordnets. Paper presented at the Proceedings of Recent Advances in Natural Language Processing, RANLP-2001, Tzigrav Chark, Bulgaria, September 5–7; pp. 149–53.
- Mendes, Sara. 2009. *Syntax and Semantics of Adjectives in Portuguese Analysis and Modelling*. Ph.D. thesis, University of Lisbon, Lisbon, Portugal.
- Mendes, Sara, and Rui Pedro Chaves. 2001. Enriching WordNet with qualia information. In *Proceedings of NAACL 2001 Workshop on WordNet and Other Lexical Resources*. Available online: <http://citeseerx.ist.psu.edu/viewdoc/download?sessionid=C77AC228C66D794A11FAAEC574329B3E49?doi=10.1.1.20.7438&rep=rep1&type=pdf> (accessed on 1 December 2023).
- Miller, George, Richard Beckwith, Christiane Fellbaum, Derek Gross, and Katherine Miller. 1990. Introduction to WordNet: An on-line lexical database. *International Journal of Lexicography* 3: 235–44. [CrossRef]

- Motos, Raquel Martínez. 2011. The Role of Interdisciplinarity in Lexicography and Lexicology. In *New Approaches to Specialized English Lexicology and Lexicography*. Newcastle upon Tyne: Cambridge Scholars Publishing, pp. 3–15.
- Pearson, Jennifer. 1998. *Terms in Context*. Amsterdam: John Benjamins Publishing Company.
- Pedersen, Ted, Serguei Pakhomov, Bridget McInnes, and Ying Liu. 2012. Measuring the similarity and relatedness of concepts in the medical domain. Paper presented at the 2nd ACM SIGHIT International Health Informatics Symposium, Miami, FL, USA, January 28–30; New York: ACM, pp. 879–80. [CrossRef]
- Pustejovsky, James. 1995. *The Generative Lexicon*. Massachusetts: MIT Press.
- Qiu, Mengyang, Nichol Castro, and Brendan Johns. 2021. Structural Comparisons of Noun and Verb Networks in the Mental Lexicon. *Proceedings of the Annual Meeting of the Cognitive Science Society* 43: 1649–55. Available online: <https://escholarship.org/uc/item/4b20s6wp> (accessed on 1 December 2023).
- Roventini, Adriana, and Rita Marinelli. 2004. Extending the Italian WordNet with the Specialized Language of the Maritime Domain. Paper presented at the 2nd International Global Wordnet Conference, Brno, Czech Republic, January 20–23; pp. 193–98. Available online: <https://www.fi.muni.cz/gwc2004/> (accessed on 1 December 2023).
- Roventini, Adriana, Antonietta Alonge, Nicoletta Calzolari, Bernardo Magnini, and Francesca Bertagna. 2000. ItalWordNet: A Large Semantic Database for Italian. Paper presented at the 2nd International Conference on Language Resources and Evaluation, Athens, Greece, May 31–June 2; pp. 783–90. Available online: <http://www.lrec-50conf.org/proceedings/lrec2000/pdf/129.pdf> (accessed on 1 December 2023).
- Sagri, Maria Teresa, Daniela Tiscornia, and Francesca Bertagna. 2004. Jur-WordNet. Paper presented at the 2nd International Global Wordnet Conference, Brno, Czech Republic, January 20–23; pp. 305–10. Available online: <https://www.fi.muni.cz/gwc2004/> (accessed on 1 December 2023).
- Siew, Cynthia. 2022. Investigating Cognitive Network Models of Learners' Knowledge Representation. *Journal of Learning Analytics* 9: 120–29. [CrossRef]
- Siew, Cynthia S. Q., and Anuta Guru. 2023. Investigating the network structure of domain-specific knowledge using the semantic fluency task. *Memory & Cognition* 51: 623–46. [CrossRef]
- Smith, Barry, and Christiane Fellbaum. 2004. Medical WordNet: A New Methodology for the Construction and Validation of Information Resources for Consumer Health. Paper presented at the COLING—The 20th International Conference on Computational Linguistics, Geneva, Switzerland, August 23–27; Geneva: Association for Computational Linguistics, pp. 371–83. Available online: <https://dl.acm.org/citation.cfm?id=1220409> (accessed on 1 December 2023).
- Vendler, Zeno. 1967. *Linguistics in Philosophy*. Ithaca: Cornell University Press.
- Vossen, Piek. 2002. *EuroWordNet General Document*. Amsterdam: University of Amsterdam. Available online: <https://research.vu.nl/ws/portalfiles/portal/77020259/EWNGeneral> (accessed on 1 December 2023).
- Vossen, Piek, Marten Postma, and Filip Ilievski. 2018. ReferenceNet: A semantic-pragmatic network for capturing reference relations. Paper presented at the 9th Global Wordnet Conference, Singapore, January 8–12; pp. 219–28. Available online: <http://compling.hss.ntu.edu.sg/events/2018-gwc/pdfs/gwc-2018-proceedings.pdf> (accessed on 1 December 2023).
- Winowiecki, Leigh, Sean Smukler, Kenneth Shirley, Roseline Remans, Gretchen Peltier, Erin Lothes, Elisabeth King, Liza Comita, Sandra Baptista, and Leontine Alkema. 2011. Tools for enhancing interdisciplinary communication. *Sustainability: Science, Practice and Policy* 7: 74–80. [CrossRef]
- Wulff, Dirk U., Simon De Deyne, Michael N. Jones, Rui Mata, and The Aging Lexicon Consortium. 2019. New perspectives on the aging lexicon. *Trends in Cognitive Sciences* 23: 686–98. [CrossRef] [PubMed]
- Yee, Eiling, Michael N. Jones, and Ken McRae. 2018. Semantic memory. In *The Stevens' Handbook of Experimental Psychology and Cognitive Neuroscience*. New York: Wiley, vol. 3. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.