

Article

A Novel Approach to Semic Analysis: Extraction of Atoms of Meaning to Study Polysemy and Polyreferentiality

Vanessa Bonato ^{1,*}, Giorgio Maria Di Nunzio ²  and Federica Vezzani ^{1,*} ¹ Department of Linguistic and Literary Studies, University of Padova, 35122 Padova, PD, Italy² Department of Information Engineering, University of Padova, 35122 Padova, PD, Italy; giorgiomaria.dinunzio@unipd.it

* Correspondence: vanessa.bonato15@gmail.com (V.B.); federica.vezzani@unipd.it (F.V.)

Abstract: Semic analysis is a linguistic technique aimed at methodically factorizing the meaning of terms into a collection of minimum non-decomposable atoms of meaning. In this study, we propose a methodology targeted at enhancing the systematicity of semic analysis of medical terminology in order to increase the quality of the creation of the set of atoms of meaning and improve the identification of concepts, as well as enhance specialized domain studies. Our approach is based on: (1) a semi-automatic domain-specific corpus-based extraction of semes, (2) the application of the property of termhood to address the diaphasic and the diastratic variations of language, (3) the automatic lemmatization of semes, and (4) seme weighting to establish the order of semes in the sememe. The paper explores the distinction between denotative and connotative semes, offering insights into polysemy and polyreferentiality in medical terminology.

Keywords: semic analysis; medical terminology; polysemy; polyreferentiality



Citation: Bonato, Vanessa, Giorgio Maria Di Nunzio, and Federica Vezzani. 2024. A Novel Approach to Semic Analysis: Extraction of Atoms of Meaning to Study Polysemy and Polyreferentiality. *Languages* 9: 121. <https://doi.org/10.3390/languages9040121>

Academic Editor: Julien Longhi

Received: 15 January 2024

Revised: 16 March 2024

Accepted: 18 March 2024

Published: 27 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/>).

1. Introduction

Semic analysis is a method of analyzing word meaning, formulated in the framework of the interpretative semantics theory developed by [Rastier \(1987\)](#). This analysis consists of the methodical factorization of the meaning of terms into semes, which are minimum non-decomposable atoms of meaning. The decomposition into atoms of meaning enables users to achieve fine-grained knowledge of the semantic dimension of language. By conducting semic analysis, the comprehension of terms can be developed and enhanced, consequently supporting a proper usage of terminology. By means of this approach, terms can be systematically and distinctively represented in the form of a sememe, which is a comprehensive collection of semes. The basic principle of the decomposition of the meaning of terms into semes can be exemplified by providing the following semic analysis of the English medical term “exophthalmic”, formulated by [Stan \(2014, p. 77\)](#). As can be observed, each /seme/ is enclosed by the slash “/” character: /regarding/ /defect/ /which affects/ /the eyes/ /shape/¹

By inference, semic analysis can also enable the identification of concepts that are linguistically designated by terms. According to the [ISO 1087: 2019 \(2019\)](#), a concept is a “unit of knowledge created by a unique combination of characteristics”. The characteristics correspond at the conceptual level to the denotative semes that are associated with the term at the linguistic level ([Vezzani et al. 2023](#)). Specifically, denotation is the neutral meaning of a linguistic form, which characterizes the conventional usage of a linguistic sign and can be found in dictionaries ([Bloomfield 1933](#)). As the denotative semes included in the sememe exclusively correspond to the unique combination of characteristics that make up a concept, we can indirectly use the results of the semic analysis to distinguish one concept from another.

In addition to denotation, semic analysis makes it possible to express the connotative meaning that terms can acquire, which belongs solely to the linguistic dimension of terminology. Indeed, the sememe also comprises connotative semes. According to Bloomfield (1933), connotation is a collection of “supplementary values” that are attached to a linguistic sign. These subjectively based values depend on cultural and social factors, as well as on linguistic influences (Bloomfield 1933). For instance, “[t]he connotation of *technical* forms gets its flavor from the standing of the trade or craft from which they are taken. Sea-terms sound ready, honest, and devil-may-care [. . .]; legal terms precise and a bit tricky [. . .]; criminals’ terms crass but to the point” (Bloomfield 1933, pp. 152–53).

The distinction between denotative and connotative semes in semic analysis opens new perspectives on the study of polysemy and polyreferentiality in terminology. In particular, in this paper we adhere to the interpretation of polysemy and polyreferentiality formulated by Vezzani et al. (2023). According to the authors, polysemy is a linguistic phenomenon occurring when a term (designating one concept) acquires different senses, which is something that also occurs because of the phenomenon of connotation. The multiplicity of senses, however, does not entail a multiplicity of concepts. This interpretation of polysemy differs from complementary notions that have been studied in terminology, such as conceptual variation (Freixa and Fernández-Silva 2017), contextual variation (León-Araúz et al. 2013), vagueness (Geeraerts 1993) and microsenses (Cruse 2001). On the other hand, polyreferentiality involves the association of multiple concepts with multiple terms that share the same lexical form (Vezzani et al. 2023). Consequently, as concepts differ from one from another, it follows that their respective terms are also different. Specifically, this interpretation of polyreferentiality closely mirrors the traditional understanding of polysemy in Linguistics.

The application of semic analysis enables the achievement of three objectives: (1) by inference, the distinction between concepts, (2) the thorough examination of the denotative and the connotative semes of terms, and (3) the analysis of polysemy and polyreferentiality of terms.

In this paper, we present a study of semic analysis of medical language. Specifically, we present an extension, development, and integration of the methodology for a systematic semic analysis of medical terminology proposed by Bonato et al. (2021). The main research goals of this paper are:

1. The proposal of a defined procedure for conducting semic analysis, with the goal of improving the systematicity of semic analysis of medical terms, and determining an approach that automates the performance of semic analysis to a greater extent.
2. The proposal of an approach allowing the analysis of polysemy and polyreferentiality in the context of semic analysis of medical terminology. Specifically, we integrate some theoretical considerations concerning the analysis of polysemy to the methodology developed by Bonato et al. (2021).

The remainder of this paper is structured as follows: in Section 2, we detail an overview of background studies related to semic analysis, with specific reference to its application in the domain of medical terminology, its performance and knowledge representation conducted therein, and the phenomenon of polysemy. In Section 3, we present the theoretical background underlying our study and we define our approach to the analysis of polysemy and polyreferentiality. Then, we describe, step-by-step, the methodology for a systematic semic analysis of medical terminology. In Section 4, we illustrate a case study, of the analysis of polysemy and polyreferentiality in semic analysis, conducted on a set of medical terms. In Section 5, we provide conclusions and suggestions for further research.

2. State-of-the-Art

In this section, we present the state-of-the-art in semic analysis and focus on the medical domain. We divide the content into three parts: the process of semic analysis, the automatization of the process and the organization of the data, and the issues related to polysemy and polyreferentiality.

2.1. Semic Analysis of Medical Terminology

Semic analysis can be proficiently adopted for the purpose of analyzing medical terminology, which is embedded in the broader domain of special languages. A special language is defined as the language used by a circumscribed set of users to accomplish the communicative needs relating to a specialized domain of knowledge (Cortelazzo 1994). Primarily, such communicative needs are referential.

Regarding special languages, Gotti (2008, p. 35) mentions the “lack of emotive connotations” as a characteristic that is emphasized “in the literature”. Moreover, the author distinguishes between words and terms: while words oftentimes possess connotative richness, “terms have a purely denotative function”. Nevertheless, “if the pragmatic purpose is persuasive [. . .], the emphasis on emotion surfaces also in specialized texts” (Gotti 2008, p. 36). Indeed, connotation is a linguistic phenomenon that also manifests in terminology belonging to specialized fields of knowledge. As considered by Bloomfield (1933, p. 152), “[i]n the case of scientific terms, we manage to keep the meaning nearly free from connotative factors, though even here we may be unsuccessful [. . .]”. For example, a negative connotation is attached to the medical term *vegetative state* (Dury 2013).

As far as the medical domain is concerned, terminology possesses connotative meaning (Dury 2013) and denotative meaning, which can both be methodically expressed in the form of semes. Therefore, the examination of both denotative and connotative semes of medical terms by semic analysis makes it possible to thoroughly analyze the semantic dimension of medical language.

Medical terms have previously been analyzed through semic analysis in several studies. For instance, Stan (2014) applied this linguistic technique with respect to terminology related to physical defects. In particular, the study aimed to “classify the semantic field of physical defects in Romanian according to specialists’ typologies and taxonomies” and, subsequently, to “identify and formulate the semic definitions of the terms in this semantic field, in relation to their lexicographic definitions” (Stan 2014, p. 65). In the study, the author extracts terms that designate physical defects from general monolingual dictionaries compiled in the Romanian language. Then, the author proposes the semic analysis of such terms, “added to and phrased based on the lexicographic definitions” (Stan 2014, p. 71). The author concludes that “the benefit of the semic analysis is that it can reveal the semantic relations between the terms, i.e., synonymy, antonymy, polysemy etc” (Stan 2014, p. 74).

Semic analysis is also provided in the medical terminological records contained in the multilingual terminological database TriMED (Vezzani and Di Nunzio 2020). In the resource, semic analysis is obtained by extracting terms from definitions found on websites, and in lexicographic dictionaries and medical terminological dictionaries. The inclusion of semic analysis in TriMED aims specifically to support medical specialized translation (Vezzani and Di Nunzio 2020, p. 278).

Moreover, Di Nunzio and Vezzani (2018) conducted a study in which the semic analysis of medical terms was used for the purpose of query rewriting. Query rewriting in information retrieval refers to the process of modifying a user’s original search query to improve the retrieval of relevant documents from a search engine. The goal of query rewriting is to enhance the effectiveness of information retrieval systems by reformulating the user’s query to better match the documents in the collection. The authors proposed a methodology aimed at replicating the experiment that was previously carried out by researchers within the framework of the CLEF eHealth Task 2: “Technologically Assisted Reviews in Empirical Medicine”. In particular, Master’s Degree in Modern Languages for International Communication and Cooperation students at the University of Padua were asked to reformulate queries previously formulated by experts. The students analyzed the terminology extracted from the queries by compiling terminological records based on the TriMED model. By using the semic analysis included in the terminological records, they formulated one of the two query variants (see the example on page 33 of Di Nunzio and Vezzani (2018)). The application of the methodology led to improved results in terms of recall and costs, compared to those obtained by CLEF participants.

Furthermore, semic analysis of medical terminology represents one of the main focuses of a contribution presented by [Bonato et al. \(2021\)](#). The authors proposed a methodology that encompasses two criteria aimed at systematizing its performance. Subsequently, the authors perform an experimental study to assess the existence of a productive interrelation between semic analysis and word embeddings. Specifically, the investigation concerned “the capability of word embeddings to retrieve semic elements in the context of medical terminology” ([Bonato et al. 2021](#), p. 223). The connection between semic analysis and word embeddings is finally identified “at the linguistic level” ([Bonato et al. 2021](#), p. 227).

In addition, semic analysis has been adopted in a paper concerning causative verbs in terminology specifically belonging to the medical domain ([Staicu 2021](#)). In particular, the author initially conducts the semic analysis of terms by adopting the perspective that involves the elaboration of multiple sememes for each lexeme, one for each single sense designated by it. Then, the author describes the terms from a syntagmatic viewpoint and, finally, delimits the subclasses of nouns with which such terms combine in the framework of medical discourse. The case study is conducted by using resources such as explanatory dictionaries, medical dictionaries, specialist treatises, medical journals and interviews with medical specialists. In particular, the author affirms that the syntactic and the lexical–semantic distribution enable the highlighting of the difference between words.

Semic analysis has also been considered in a study authored by [Vezzani \(2023\)](#), which concerned the investigation of the connotation of somatic terminology in the field of oncology in the Italian and French languages. The paper focuses on the comparative analysis of the pairs of terms *seno/sein* and *mammella/mamelle*, with the aim of identifying the connotative charge that medical terms can acquire. The linguistic analysis conducted by the author relies on data from different sources, such as Sketch Engine, medical websites, terminological resources, the online encyclopedia Wikipedia, lexicographic dictionaries and medical dictionaries. In the context of the article, the manually conducted semic analysis helps to identify the diverging connotation ascribed to the pair of terms *mammella/mamelle*. The examination of the connotation of medical terms emerges as fundamental to the avoidance of translation errors.

2.2. Knowledge Representation and Semi-Automatic Semic Analysis

Multiple studies have focused on the performance and the representation of semic analysis, in terms of computer data structures. Between the decade of the 1990s and the first decade of the 2000s, two of the most important resources that changed the panorama of the management and representation of language resources were WordNet ([Miller 1995](#)) and HowNet ([Dong and Dong 2003, 2006](#)). While these two resources share the same idea of organizing linguistic knowledge, their view of their semantic organization is different. This is well expressed by [Veale \(2005, p. 1149\)](#): “WordNet is *differential* in nature: rather than attempting to express the meaning of a word explicitly, WordNet instead differentiates words with different meanings by placing them in different synonym sets, and further differentiates these synsets from one another by assigning them to different positions in its taxonomy. In contrast, HowNet is *constructive* in nature. It does not provide a human-oriented textual gloss for each lexical concept, but instead combines sememes from a less discriminating taxonomy to compose a semantic representation of meaning for each word sense.” Besides their semantic organization, two of the most important differences between WordNet and HowNet are one is based on words and synsets (a set of synonyms that share the same meaning), while the other is based on concepts and sememes. This is also why some authors mention HowNet as a sememe knowledge base (SKB), in which “every word [...] contains some senses, and each sense is annotated with a hierarchical structure of sememes, i.e., sememe tree” ([Du et al. 2020](#)).

Given the nature of our work, in this section we will focus on HowNet and how this resource has been used in numerous studies concerning sememes ([Niu et al. 2017](#); [Li et al. 2018](#); [Liu et al. 2020](#)). In particular, we will focus on those contributions that deal with the topics of sememe prediction and the construction of SKB.

Sememe prediction is an activity conceived by Xie et al. (2017) that is targeted at suggesting sememes in an automatic mode for words that lack this kind of annotation (Du et al. 2020). The fundamental principle at the basis of sememe prediction is that “overlapped sememes” may be shared by words that are characterized by similarity, in terms of their “semantic meanings” (Xie et al. 2017, p. 4200). In this sense, the most challenging issue that the task involves is “how to represent semantic meanings of words and sememes to model the semantic relatedness between them” (Xie et al. 2017, p. 4200). For instance, Xie et al. (2017) propose three different models to perform it, namely sememe prediction with word embeddings (SPWE), with sememe embeddings (SPSE) and with aggregated sememe embeddings (SPASE). Li et al. (2018) present a contribution in which word descriptions found on wiki websites are used to perform the prediction of lexical sememes in an automatic way. A different approach to sememe prediction is presented in the study by Du et al. (2020), in which definitions from dictionaries are used to accomplish the task. Furthermore, a diverse perspective is presented in the contribution of Ye et al. (2022, p. 128), which focuses on conferring a structured organization to sememes, with the specific purpose of “predicting a sememe tree with hierarchical structure rather than a set of sememes”. In addition, Qin et al. (2023) tackle the activity of sememe knowledge modeling by utilizing neural networks. Moreover, in a recent study by Zhang et al. (2023, p. 2790), an innovative method aimed at “aligning BabelNet synsets to HowNet senses” is proposed. Specifically, BabelNet (Navigli and Ponzetto 2012) is an encyclopedic dictionary whose nature is multilingual.

Several studies concerning the construction of SKBs have also been carried out. Qi et al. (2021) pioneered the formulation of a method involving the utilization of dictionaries for sememe knowledge base creation. The authors applied this entirely automatic method to create two sememe knowledge bases, respectively concerning the English and French languages. The issue regarding the building of SKBs was also tackled from a multilingual perspective. Indeed, Qi et al. (2022) created a sememe knowledge base characterized by being multilingual, which was built on the BabelNet resource. More particularly, the authors perform sememe prediction for BabelNet synsets (SPBS) by means of an automatic approach, so that “the words in many languages in the synset would obtain sememe annotations simultaneously” (Qi et al. 2022, p. 158). To do so, the authors encode multiple kind of information extracted from BabelNet in their model, with specific reference to multilingual synonyms and glosses, as well as information related to semantics obtained from images.

Semic analysis has also been tested for its capability to detect “text semantic content” (Grzesitchak et al. 2007, p. 227). In particular, the authors developed the software SEMY, whose purpose is to provide “a semi-automatic semantic profile of texts or linguistic entities” (Baider 2014). As detailed by Baider (2009, p. 33), it is “based on Rastier’s semic feature theory” and makes it possible to identify textual isotopies (Rastier 1987). Grzesitchak et al. (2007, p. 229) formulated the assumption according to which “a definition is a sememe in textual form”. Indeed, as specified by the authors, word extraction is performed from dictionary definitions that are specifically contained in the lexicographic resource *Trésor de la Langue Française Informatisé* (Dendien and Pierrel 2003). As Grzesitchak et al. (2007, p. 229) indicate, “les mots pleins d’une définition (substantifs, adjectifs, verbes et certains adverbes) sont, une fois lemmatisés, considérés comme les sèmes potentiels qui constituent le signifié d’une unité lexicale”. SEMY has been used in multiple studies (Baider 2009; Reutenauer et al. 2009; Baider 2014), including in a contribution targeted at the identification of semes for the French word *Outreau* (Reutenauer et al. 2010).

The issue of the identification of semes was addressed by Peláez and Mateos (2018, p. 70), who proposed a methodology involving the use of a semantic formula, defined as “a lexeme’s group of denoted and connoted semantic classes”. The semantic formula is a “kind of semantic framework”, and it is “formulated in terms of the semantic classes to which the particular entities, attributes, actions, states, processes and relations etc. belong and which a lexeme calls to mind (i.e., which it denotes/connotes)” (Peláez and Mateos

2018, pp. XXXII–XXXIII). In particular, the formula provides a “pattern” that enables the development of semic analysis which takes as a basis the detection of five semantic classes: “events, entities, attributes, relations, and determinations” (Peláez and Mateos 2018, p. 103). Then, “[o]nce the semantic formula is established, the lexeme’s generic and specific nuclear semes can be determined (i.e., its semic development)” (Peláez and Mateos 2018, p. 116). Specifically, generic semes “transport fundamental aspects of meaning related especially to entities (number, gender) and to events (mode, tense, aspect, voice)”. Specific semes instead, “concern those semantic features necessary to identify the concept referred to by a lexeme [...]. They are identified by comparing the particular lexeme [...] to lexemes belonging to the same semantic domain, thereby establishing both common and distinctive semantic features”. Contextual semes, which express the meaning that a lexeme assumes in context, are then added to the nuclear ones.

Regarding the performance of semic analysis, Pottier (1992, p. 117) stated that “[w]hat is surprising is the arbitrariness of the choice of semes as compared to the perceptible world”. Taking this assertion as a starting point, Bonato et al. (2021) developed a methodology aimed at improving the systematization of semic analysis of medical terminology and reducing subjective manifestations in its application. Specifically, the methodology presents two main criteria. The first criterion consists of the adoption of a domain-specific corpus-based approach. In particular, semes are extracted from a corpus of specialized intensional definitions, and the mesogeneric seme is then added to the generated sememe. Specifically, the mesogeneric seme expresses belonging to a specific domain (Prié 1995). As defined by Roche (2012, p. 26), the “[i]ntensional definition [...] comprises the *superordinate concept* immediately above followed by one or several *delimiting characteristics*” (Roche 2012, p. 26). The authors used definitions extracted from two medical dictionaries: the Merriam–Webster Medical Dictionary² and the TheFreeDictionary’s Medical Dictionary³. The second criterion that composes the methodology concerns the application of the property of termhood (Kageura and Umino 1996) to convey both the diastratic and the diaphasic variations of language. According to Bonato et al. (2021, p. 222), “[t]he interconnection of the principles of termhood, diastratic and diaphasic dimensions would improve the specificity which could characterize the representation of terminology from a conceptual viewpoint. Indeed, the domain-contextual knowledge would be captured fostering the adequate achievement of a distinction between medical technical and non-technical terms”. Notwithstanding the adoption of this approach, subjectively biased phenomena could still be expected to occur, hence determining the necessity to further develop the methodology.

2.3. The Phenomenon of Polysemy

The realization of a thorough semic representation, however, also implies the necessity to consider and manage the phenomenon of polysemy. Numerous studies in terminology, specialized lexicography, and lexicography focus on the topic of polysemy (Frassi 2022; León-Araúz and San Martín 2011; L’Homme 2020b, 2023; Pethő 2001; Sterner 2022; Vezzani et al. 2023; Rastier 2014; Condamines and Rebeyrolle 1997). However, the state-of-the-art lacks a general consensus as to its concept, consequently determining the multiplicity of its theoretical definitions. For example, in the words of Corbeil (1988, p. 57), polysemy occurs when “la même dénomination semble correspondre à des notions différentes, en tout ou en partie (polysémie terminologique)”. According to Condamines and Rebeyrolle (1997, p. 174), polysemy manifests when “à un signifiant unique correspondent plusieurs sens perçus comme étant reliés entre eux”.

In particular, it is relevant to notice that the notions of polysemy and synonymy in terminology vary depending on the different interpretations given by different scholars. For instance, in the view of Rouleau (2003, p. 146), “[i]déalement, le rapport entre un terme et sa notion devrait, en langue de spécialité, être monosémique (un terme par notion). Dans la pratique, il n’en est pas toujours ainsi. Il arrive qu’un même terme recouvre plusieurs notions (polysémie) ou encore que plusieurs termes désignent une même notion (synonymie)”. According to L’Homme (2020a, p. 105), polysemy is “a phenomenon in

which the same lexical form has multiple meanings”, and “polysemous items usually share at least one semantic component”. Specifically, the author affirms that polysemy can be frequently encountered in specialized fields of knowledge and may emerge within one single domain (L’Homme 2023). In the words of Buysschaert (2021, p. 67), polysemy “occurs when a given term has more than one meaning and the meanings are related”. The author, who also refers to the framework of special languages, states that the avoidance of polysemy should occur whenever feasible. Despite this, the phenomenon affects medical terminology (Rouleau 2003; Džuganová 2017; Buysschaert 2021). Indeed, in the view of Buysschaert (2021, p. 67), an exemplification of polysemy can be represented by the term *drug*, which can be used to designate both “a therapeutic agent” and “a stimulating or depressing substance that is potentially addictive”. As can be observed, there are different interpretations of the concept of polysemy. In this study, however, we adhere to the interpretation of the concept of polysemy that was proposed by Vezzani et al. (2023), which differs from the above-mentioned theoretical perspectives proposed by scholars (see Section 3.1).

2.4. Polysemy in Semic Analysis

The phenomenon of polysemy has previously been investigated in the context of semic analysis (Cusimano 2007; Thomas 2017, 2023; Conso 2020). As outlined and listed by Cusimano (2007), it is possible to identify three diverging modes of operation in the management of polysemy in semic analysis: (1) the definition of a “sémème d’un lexème polysémique par l’intersection sémique de ces différentes significations” (Touratier 2000), (2) the presence of different sememes, as in the view of Picoche (1986) and Rastier (1987), and (3) the “tentatives de contournement du problème du nombre de sémèmes qu’engage la polysémie” (Cusimano 2007, p. 92), represented by theories such as contextualism, constructivism, the theory of cognition and the prototype theory.

3. A Methodology for a Systematic Semic Analysis of Medical Terminology

In this section, we present the methodology for a systematic semic analysis of medical terminology. We start with the description of the theoretical background of the work in Section 3.1. Subsequently, in Section 3.2, we define the approach that we adopt to the analysis of polysemy and polyreferentiality. Finally, in Section 3.3, we present the procedure for the systematic semic analysis.

3.1. Theoretical Background on Polysemy and Polyreferentiality

As we previously mentioned, semic analysis enables the analysis of terms from a linguistic viewpoint. In this paper, we start from the assumption that semic analysis can also be used to analyze the conceptual dimension of terminology. This is based on the idea that the unique combination of denotative semes that compose the sememe of a term corresponds to the unique combination of characteristics that make up the concept.

The consideration of these two fundamental dimensions characterizes the theoretical approach that we adopt in this contribution, which specifically assumes that the discipline of terminology features the indispensable interrelation of a dual dimension: conceptual and linguistic (Costa 2013; Santos and Costa 2015; Carvalho et al. 2016). Notwithstanding their complementarity, these two dimensions are inherently dissimilar: while the conceptual dimension focuses on the concept, the linguistic dimension concerns the term that designates the concept. In particular, the view that we adopt on the concept of *concept* differs from the one embraced by Rastier (2015), who stated that “[u]n concept est un sémème construit, dont la définition est stabilisée par les normes d’une discipline, de telle façon que ses occurrences soient identiques à son type”. In our perspective, indeed, the concept is not a sememe, as the sememe exclusively concerns the linguistic dimension of terminology.

Semic analysis also makes it possible to express both the denotative and the connotative meaning of terms. Indeed, terms can additionally acquire connotation due to specific contextual usage and to the semantic evolution that may affect language over the course of

time. In particular, denotation and connotation are respectively represented by inherent semes and afferent semes. Specifically, the sememe, defined in Section 1 as a comprehensive collection of semes, comprises both “inherent semes” and “afferent semes” (Rastier 1987, 2005). An inherent seme is a seme belonging to a sememe’s type, while afferent semes solely occur in the sememe that represents the token, as their presence exclusively manifests when “contextual indication” occurs (Hébert 2020). In this contribution, we hereafter use the term “denotative semes” to refer to inherent semes, and the term “connotative semes” to refer to afferent semes.

Since semic analysis allows the distinguishing of denotative semes and connotative semes, it also makes it possible to analyze polysemy and polyreferentiality in terminology. In particular, by embracing the approach formulated by Vezzani et al. (2023), we consider that the characteristics that make up the concept precisely correspond to the semes that express the denotation of the term. Polysemy occurs when a term acquires connotation in a specific context of usage, thus assuming a different sense. Hence, polysemy is a linguistic phenomenon that concerns the multiplicity of senses that a term (designating one concept) can acquire. As we already mentioned, the multiplicity of senses does not entail the multiplicity of concepts. By taking this standpoint, it is possible to draw a fundamental distinction between polysemy and polyreferentiality. Indeed, polyreferentiality is a phenomenon related to the conceptual dimension of terms, which takes place when multiple concepts are associated with a single lexical form. These multiple concepts, however, solely share the same lexical form, which is to say the same sequence of characters. Consequently, as concepts differ one from another, it follows that their respective terms are also different.

In the following section, we illustrate the methodological approach that we adopt to analyze polysemy and polyreferentiality in semic analysis.

3.2. Methodological Approach

As we previously mentioned, terms can contextually acquire connotation. More generally, certain terms may lose or gain their connotation in specific contexts due to various factors, such as semantic shift and cultural changes. For example, “queer” was primarily used as a pejorative term to insult or denigrate individuals who were perceived as non-conforming, with regard to traditional gender and sexual norms. Today, “queer” is often used as an umbrella term to encompass a wide range of non-heteronormative sexual orientations and gender identities. By considering this, it is possible to infer that connotation does not occur in the totality of the contexts of a term’s usage. It follows that the absence of connotation implies the existence of a sememe that is uniquely composed by denotative semes. In our perspective, this sememe expresses the “neutral meaning” that a term possesses. In this sense, we consider that the “neutral meaning” of a term is entirely devoid of connotation. Conversely, the presence of connotation determines the existence of sememes that include both the semes that represent the denotation and the semes that convey the connotation of a term. In this case, the sememe consists of both denotative and connotative semes.

In particular, according to Vezzani et al. (2023), the ascription of connotation to a term determines the occurrence of the linguistic phenomenon of polysemy, which involves the variation (and multiplicity) of senses that affects the term. By adopting this stance, we consider that each distinct sememe that contains connotative semes corresponds to one of the senses acquired by a term, therefore expressing the polysemous nature of the term. At this point, it is also possible to identify the conditions required for the different senses of a term to differ one from another: (1) the sharing of the same denotative semes, and (2) the diversity in at least one connotative seme. Specifically, the multiple senses of a term are different at the linguistic level. On the other hand, polyreferentiality entails the reference to different concepts. Accordingly, we assume that terms that uniquely share the same sequence of characters but designate different concepts are represented at the semic level by different sememes. Indeed, these different concepts are made up of a different and unique combination of characteristics. Consequently, their sememes are composed

by a different and unique combination of semes. For this reason, the single condition that must be satisfied for these terms to differ from one another is the diversity in at least one denotative seme.

By adopting this approach, we aim at providing a defined stance on the analysis of polysemy and polyreferentiality in the framework of the proposed methodology. Moreover, we expect to reach toward a higher systematicity in the performance of semic analysis, as additional procedural criteria are embedded in the methodology.

3.3. A Procedure for a Systematic Semic Analysis of Medical Terminology

Based on the methodology for a systematic semic analysis of medical terminology proposed by [Bonato et al. \(2021\)](#), we hereafter list the revisited procedural steps that compose the updated methodology:

1. Semi-automatic domain-specific corpus-based extraction of semes;
2. Termhood to address the diaphasic and the diastratic variations of language;
3. Automatic lemmatization of semes;
4. Seme weighting to establish the order of semes in the sememe.

In particular, we update the methodology presented by [Bonato et al. \(2021\)](#) by proposing a semi-automatic domain-specific corpus-based extraction of semes, instead of a manual extraction. Moreover, we add two additional steps to the methodology: the automatic lemmatization of semes and seme weighting to order semes. The purpose of the application of these sequential phases is to improve the systematicity and the automaticity of the performance of semic analysis. Indeed, one of the goals of our study is to effectively minimize subjective manifestations in its performance. These steps cover the multiple actions involved in semic analysis, ranging from the origination of the sememe to its final composition.

3.3.1. Semi-Automatic Domain-Specific Corpus-Based Extraction of Semes

The first criterion comprised in the methodology presented by [Bonato et al. \(2021\)](#) involves the creation of a corpus of specialized intensional definitions, from which semes are extracted. Furthermore, the mesogeneric seme that indicates the domain of a term is additionally included in the sememe.

In this paper, we specifically propose to semi-automatically extract candidate semes from specialized intensional definitions of terms. In particular, the corpus is manually compiled and our approach starts with the recollection of definitions from the Merriam–Webster Medical Dictionary and TheFreeDictionary’s Medical Dictionary, which are domain-oriented dictionaries. The definitions included in TheFreeDictionary’s Medical Dictionary that constitute our corpus are extracted from the Miller–Keane Encyclopedia and Dictionary of Medicine, Nursing, and Allied Health, Seventh Edition, the Segen’s Medical Dictionary, the Farlex Partner Medical Dictionary, the American Heritage Medical Dictionary, the Medical Dictionary for the Health Professions and Nursing, the Collins Dictionary of Medicine, the Medical Dictionary for the Dental Professions, the Collins Dictionary of Biology, 3rd edition, and the McGraw-Hill Concise Dictionary of Modern Medicine. We additionally resort to general language dictionaries, also to collect information concerning the connotation of terms. Then, we reformulate these definitions to generate intensional definitions. However, we do not include hyponyms in the reformulated definitions of the terms.

For example, we describe the process of reformulation by considering one of the definitions of the reference term *capsule* comprised in the Merriam–Webster Medical Dictionary: We:

A viscous or gelatinous often polysaccharide envelope surrounding certain microscopic organisms (as the pneumococcus).

The process concerns the identification of the elements of the definition that must be included in the reformulated intensional definition. According to the [ISO 1087: 2019 \(2019\)](#), an intensional definition is a “*definition that conveys the intension of a concept by*

stating the immediate *generic concept* and the *delimiting characteristic(s)*". Hence, in the first place we identify the immediate generic concept that is contained in the definition, which is represented by the term *envelope*. Then, we detect the following delimiting characteristics: *viscous or gelatinous, often polysaccharide, surrounding certain microscopic organisms (as the pneumococcus)*.

The automatic extraction of medical terms from the corpus can be performed by adopting the approach for the automatic extraction of terms used in the study by [Vezzani and Di Nunzio \(2019\)](#). This approach consists of finding and extracting only those terms that match a list of medical terms. In the current work, we use the *tidytext* R package for text analyzes⁴ to extract the list of the terms that are in the corpus and match them with the list of medical terms provided by the Medical Subject Headings (MeSH) database.⁵ For example, the tool MeSH on Demand⁶ can be used to automatically match the medical terms that are extracted from the intensional definitions. In our study, the extracted terms are considered as the candidate semes that form the sememe of a medical term. When the automatic process is applied, the candidate semes that compose the sememe are automatically extracted.

Moreover, we find it necessary to precisely circumscribe the term in its specific context of usage. To this end, we maintain the inclusion of the mesogeneric seme within the sememe. As we mentioned, the mesogeneric seme indicates the domain ([Prié 1995](#)). According to [Prié \(1995, p. 17\)](#), who lists the classes of sememes on the basis of the work of Rastier, the domain is a "clustering of taxemes" that is also targeted at avoiding polysemy from a lexical viewpoint. Specifically, as [Prié \(1995\)](#) states, polysemy stems from the existence of several domains. However, as we previously pointed out, polysemy also manifests in the framework of a single domain ([L'Homme 2023](#)), and its occurrence can be observed even in medical language. According to [Bonato et al. \(2021, p. 221\)](#), the identification of mesogeneric semes "can be considered as strategically fundamental to reduce the occurrence of polysemy". For instance, the inclusion of the mesogeneric seme */medicine/* in the sememe of medical terms has been adopted in the methodology proposed by [Bonato et al. \(2021\)](#). However, the identification of the mesogeneric seme and its inclusion in the sememe is not sufficient to effectively minimize polysemy in semic representation. For this reason, in addition to the mesogeneric seme, we also include in the sememe a seme that indicates the subdomain in which the medical term is used.

The semi-automatic extraction of the semes will reduce subjectively driven manifestations in the performance of semic analysis. Indeed, we formulate intensional definitions following the criteria established by the [ISO 1087: 2019 \(2019\)](#). This means that semes are systematically extracted from definitions that are structured according to standardized criteria. The selection of semes, therefore, will rely only partially on the personal knowledge that the users of semic analysis have of the semantic dimension of medical terminology. As a consequence, even when the user does not have a high level of specialized medical knowledge, the chances of not identifying terms that belong to the medical domain can be lowered. Our structured methodology, however, does not totally exclude the occurrence of subjective manifestations. Indeed, the proposed methodology relies on available resources and dictionaries created by experts who may have made subjective choices in constructing these resources. For example, the MeSH database was compiled by experts who selected the terms that are included in it. Furthermore, equal consideration should be given to the possibility of errors in term extraction, which could lead to the risk of excluding relevant terms during the extraction process.

3.3.2. Termhood to Address the Diaphasic and the Diastratic Variations of Language

The second criterion adopted in the methodology formulated by [Bonato et al. \(2021\)](#) is the application of the property of termhood ([Kageura and Umino 1996](#)) to convey both the diastratic and the diaphasic dimensions of language.

Specifically, the diastratic dimension constitutes the "variation across socio-economic classes and social groups" ([Berruto 2010, p. 227](#)), while the diaphasic one represents the

“variation across situations”. To address these dimensions of language in the context of semic analysis, the principle of termhood can be proficiently adopted. In the words of [Kageura and Umino \(1996, pp. 260–61\)](#), termhood is defined as “the degree that a linguistic unit is related to (or more straightforwardly, represents) domain-specific concepts”.

In the methodology proposed by [Bonato et al. \(2021\)](#), termhood is expressed by contextually utilizing either specialized terms or popular ones as semes in the formulation of the sememe. Popular terms are layperson terms, which can be understood by non-experts in the medical field. More particularly, specialized terminology composes the semic analysis of a medical term when the term is specialized. Conversely, the semic analysis of popular terms is accordingly composed by popular terms. For example, within the sememe of a specialized medical term, the popular medical term *pneumoniae* is replaced with the corresponding specialized term *streptococcus pneumoniae*.

This criterion is maintained in the methodology. In this study, we rely on two resources for the identification of specialized and popular terms: (1) the web application developed by [Di Nunzio and Vezzani \(2021\)](#), which contains medical terminological data deriving from the “Multilingual Lemma Collection” website, and (2) the multilingual medical terminological database TriMED ([Vezzani and Di Nunzio 2020](#)). In particular, in the web application, each terminological record includes the distinction between the technical form of a term and its popular form. The database TriMED instead presents the differentiation between the “common name” of a term and its “scientific name”. By examining these sections, it is possible to identify the specialized terms and the popular terms contained in definitions.

3.3.3. Automatic Lemmatization of Semes

The performance of semic analysis also involves a reflection on the lexical form that semes assume. Consequently, this aspect should be also considered when developing a focused methodology. Indeed, semes may be represented by terms that present multiple inflectional endings. By performing lemmatization, we aim to ensure consistency in the lexical representation of terms in semic analysis. Indeed, all the formulations of semic analysis will exclusively contain lemmatized terms, therefore excluding subjectively based manifestations, such as the inclusion in the sememes of both lemmas and inflected forms of terms. Once more, the intended purpose is to minimize subjectivity in the selection of semes, which could concern the lexical features of terms as well.

Since medical terminology is our field of study, the automatic lemmatization of semes can be performed by utilizing BioLemmatizer ([Liu et al. 2012](#)). Specifically, BioLemmatizer is a lemmatizer purposefully developed for usage in the biomedical domain.

3.3.4. Seme Weighting to Establish the Order of Semes in the Sememe

The issue concerning the order of semes within the sememe has been discussed in the literature. In particular, different perspectives on the topic are presented by authors. For example, according to [Bianca and Piccari \(2007, p. 18\)](#), “a lexeme’s collection of semes are not arranged in sequential order”. Indeed, the authors affirm that “the meaning of a lexeme” can solely be modified by either eliminating or inserting semes ([Bianca and Piccari 2007, p. 18](#)). Conversely, [Filipec \(1994, p. 170\)](#) states that “[t]he order of semes and their hierarchy are not arbitrary; they are related to the occurrence of semes in the respective partial subsystem”. In the contribution by [Peláez and Mateos \(2018, p. 117\)](#), both “generic and specific nuclear semes” are arranged hierarchically on the basis of multiple criteria, namely “implication, presupposition, and consequence”. According to the authors, “[t]he meaning of a lexeme does not depend solely on the semes (considered independently) that comprise it, nor solely on their particular cluster (their semic nucleus), but also on the particular manner in which the nuclear semes are organized (nuclear configuration” ([Peláez and Mateos 2018, p. 64](#)).

With specific reference to the topic of semes weighting, [Reutenauer \(2012\)](#) proposes to assign a specific weight to semes found in definitions. By doing this, different levels

of importance are assigned to semes (Reutenauer 2012). In particular, Reutenauer (2012) mentions three different approaches for seme weighting: (1) the consideration of the location of semes within the definition; (2) the evaluation of the relevance of the definitions in which semes are embedded; and (3) the calculation of both occurrences of a seme, with respect to the lexicographic entry contained in dictionaries, and the total number of occurrences of such semes in the context of the entire dictionary, which can be computed by using the tf-idf method.

In our methodology, the ordering of semes by seme weighting is specifically targeted at increasing the systematicity of semic analysis. Moreover, the ordering of semes is also aimed at better representing the semantic content conveyed by a term. However, we adopt a different method for the attribution of weight to semes. More particularly, weight can be attributed to semes by applying the *v-tech* value, formulated by Vezzani (2019). Specifically, the *v-tech* value is an assessment parameter that enables the evaluation of the technicality of terms. The main idea underlying the conceptualization of the *v-tech* value is that the technical value of a term is inversely proportional to the polysemous nature of terms. In detail, the approach adopted by the author is based on the conception of the notion of “technicality” as a semantic intrinsic property of a term. In this sense, the value can be defined as corpus independent. The concept underlying the *v-tech* value therefore differs from the concept of termhood (Kageura and Umino 1996). Indeed, “le degré de termicité d’un terme [. . .] repose, en général, sur la fréquence d’apparition d’un terme candidat dans le corpus analysé” (Vezzani 2019, pp. 216–17). The topic of term technicality has also been studied by Bertels (2011). In our perspective, however, the application of the *v-tech* formula can be conceived as both pertinent and profitable in the context of semic analysis.

To sum up, the proposed methodology for a systematic semic analysis of medical terminology can be viewed as consisting of a twofold approach: corpus-based and corpus independent. In this framework, these opposing perspectives are complementarily combined and sequentially addressed, aiming to systematically represent and analyze the semantic dimension of terms.

4. A Novel Approach to the Analysis of Polysemy in Semic Analysis: A Case Study

In this section, we present a preliminary case study concerning the analysis of polysemy and polyreferentiality within the framework of the proposed methodology for a systematic semic analysis of medical terminology. We consider the first three sequential steps included in the methodology, namely: (1) the semi-automatic domain-specific corpus-based extraction of semes, (2) the application of the property of termhood to address the diaphasic and the diastratic variations of language, and (3) the automatic lemmatization of semes. Specifically, in the context of the case study, we manually apply these steps to perform the semic analysis of medical terms.

In particular, we perform the semic analysis of four medical terms whose terminological records are included in the multilingual terminological database TriMED (Vezzani and Di Nunzio 2020): *Down syndrome*, *tuberous sclerosis*, *capsule*, and *resistance*. We divide the analysis into two parts: those terms that may be polysemic, and those terms that may be polyreferential. In Section 4.1, we propose the semic analysis of the terms *Down syndrome* and *tuberous sclerosis*. Subsequently, in Section 4.2, we apply the same procedure with respect to the terms *capsule* and *resistance*.

4.1. Analysis of Polysemy

In this section, we perform the semic analysis of the terms *Down syndrome* and *tuberous sclerosis*.

We present the manually performed semic analysis of the terms resulting from the manual application of the first three steps that compose the proposed methodology. Then, we automatically match all the medical terms contained in the intensional definitions—that serve as a basis for the semic analysis—with the list of medical terms provided by the MeSH database. We use the tool MeSH on Demand for this purpose. Subsequently, we

compare the list of the MeSH medical terms automatically extracted from each intensional definition to the respective list of semes that compose the semic analysis. This comparative procedure allows us to evaluate the extent to which the automatic extraction of MeSH terms in definitions enables an accurate automatic identification of all the terms that constitute the semes of the reference terms.

4.1.1. Down Syndrome

Our analysis started by taking as its basis the intensional definitions of the term *Down syndrome*, which were derived from the reformulation of the definitions contained in the Merriam–Webster Medical Dictionary, in the TheFreeDictionary’s Medical Dictionary and in the Oxford Advanced Learner’s Dictionary⁷. We consulted these dictionaries to collect definitions that enabled us to obtain information about the contextual usages of the terms *Down syndrome* and *tuberous sclerosis*. Indeed, this kind of information is essential to detect the connotation that terms can contextually acquire. Due to the fact that information about connotation is generally found in lexicographic resources, we additionally included in our corpus the definition found in the general language dictionary Oxford Advanced Learner’s Dictionary. In Appendix A, we provide a list of all the formulated intensional definitions of the term *Down syndrome*. Following the creation of the intensional definitions, we performed the semic analysis by manually applying the first three sequential phases of the methodology. Each definition in our corpus respectively corresponds to one formulation of semic analysis. The listed sememes represent the final result of the analysis:

1. /medicine/ /pathology/ /condition/ /chromosome 21/
2. /medicine/ /pathology/ /condition/ /chromosome 21/ /mental retardation/
3. /medicine/ /pathology/ /condition/ /chromosome 21/ /learning disability/

To the generated sememes, the mesogeneric seme */medicine/*, which indicates the domain, and the seme */pathology/*, which represents the subdomain, were added.

From the examination of the formulations of semic analysis, the existence of a sememe consisting solely of denotative semes emerged—the first sememe—and two sememes—the second and the third sememes—that additionally contain connotative semes. The three sememes share the same denotative semes */condition/* and */chromosome 21/*. As we previously mentioned, when connotation is attached to a term, the linguistic phenomenon of polysemy takes place. Specifically, polysemy occurs when multiple sememes are associated with a term: (1) a sememe that is exclusively composed by denotative semes, which expresses the “neutral meaning” of the term, and (2) at least one sememe that includes both denotative semes and connotative semes. The additional inclusion in the sememe of connotative semes results in the generation of different unique combination of semes, therefore creating a different sememe. Such different sememes represent one of the senses that the term can contextually acquire. Considering that multiple sememes are associated with the term, it is possible to affirm that the term *Down syndrome* is polysemic. Specifically, the connotative semes that can be identified are */mental retardation/* and */learning disability/*.

In addition, the examination of the connotative semes that compose the sememe makes it possible to evaluate whether a positive or a negative connotation is attributed to the reference term. For instance, the occurrence of the seme */mental retardation/* in a sememe is indicative of the fact that a pejorative connotation is present. In the words of [Laureno \(2017, p. 217\)](#), “[t]erms are changed for social and political reasons. [...] With time “mental retardation” itself came to carry a negative connotation to the lay person”. Considering this, it is possible to affirm that a negative connotation is attached to the term *Down syndrome*. Furthermore, a pejorative connotation is also attributed to the term *learning disability*. Indeed, [Cluley et al. \(2022\)](#) stated that this term “has undergone much revision and critique, being linked to stigma and prejudice”.

Following the analysis of the connotative semes associated with the reference term *Down syndrome*, we automatically matched the medical terms contained in the three considered intensional definitions with the list of terms contained in the MeSH database. For each definition, we list the automatically matched terms:

- Definition 1: Trisomy; Muscle Hypotonia; Chromosomes, Human; Cognition
- Definition 2: Down Syndrome; Intellectual Disability; Chromosomes, Human; Pair 21
- Definition 3: Learning Disabilities; Chromosomes

The comparison between the automatically matched MeSH terms and the terms that compose, at the semic level, the different formulations of semic analysis revealed that the term-matching process did not enable us to identify all the terms that compose the manually generated semic analysis. As we expected, the terms that represent, at the semic level, both the mesogeneric seme and the seme that expresses the subdomain were not included in the list of extracted MeSH Terms. Indeed, both the mesogeneric seme and the seme that indicates the subdomain were not included in the generated intensional definitions. In particular, with reference to the MeSH terms that were extracted in the first intensional definition, it can be observed that the term *Chromosomes, Human* does not match the seme */chromosome 21/* from a lexical viewpoint. Moreover, the list of extracted MeSH Terms comprises terms that were not included in the corresponding formulation of semic analysis. For what concerns the second intensional definition, the negatively connoted term *Intellectual Disability* was extracted as a MeSH term. However, the term *Intellectual Disability* does not match the seme */mental retardation/*. Also in this circumstance, however, the MeSH term *Chromosomes, Human, Pair 21* does not match the seme */chromosome 21/* that is included in the sememe. Furthermore, the MeSH term *Down Syndrome* does not constitute a seme in the corresponding sememe. Regarding the examination of the list of MeSH terms that were extracted in the third intensional definition, it led to the consideration that the pejoratively connoted term *Learning Disability* was also automatically matched.

4.1.2. Tuberous Sclerosis

Subsequently, we manually applied the proposed methodology to perform the following semic analysis of the reference term *tuberous sclerosis*. We took as a basis the intensional definitions that derived from the reformulation of the definitions contained in the Merriam–Webster Medical Dictionary, in the TheFreeDictionary’s Medical Dictionary and in the online general language dictionary YourDictionary⁸. In Appendix B, we provide the list of the generated intensional definitions of the term *tuberous sclerosis*. Also in this case, each formulated intensional definition respectively corresponds to one of the listed formulations of semic analysis:

1. */medicine/ /pathology/ /brain/ /skin/*
2. */medicine/ /pathology/ /brain/ /skin/ /mental retardation/*
3. */medicine/ /pathology/ /brain/ /skin/ /mutation/*
4. */medicine/ /pathology/ /skin/ /brain/ /mutation/ /intellectual disability/*

As can be observed, the sememes of the reference term are formed by specialized medical terms. Indeed, the reference term *tuberous sclerosis* belongs to the specialized linguistic register. After the formulation of the sememes, we added the mesogeneric seme */medicine/* that specifies the domain of usage of the reference term and the seme */pathology/* that indicates the medical subdomain.

The examination of the different formulations of semic analysis of the reference term *tuberous sclerosis* led to the identification of a sememe that is exclusively constituted by denotative semes and three different sememes in which both denotative semes and connotative semes are present. Specifically, the four sememes share the same denotative semes */brain/* and */skin/*. The association of multiple sememes with the term is indicative of the fact that the term *tuberous sclerosis* is polysemic. In particular, the connotative seme */mental retardation/* that is contained in the sememe determines that a negative connotation is attached to the reference term. Moreover, the seme */intellectual disability/* also carries a negative connotative charge. As a matter of fact, as affirmed by Snipstad (2022, p. 109), “[i]n the diagnostic manuals, intellectual disability appears to be associated with some less than positive connotation”. Furthermore, the seme */mutation/* also presents a pejorative connotative charge. Specifically, Jarvik and Evans (2017, p. 491) highlight “the popular conflation of “mutation” with the grotesque and disturbing”.

Following the analysis of the connotative semes of the reference term *tuberous sclerosis*, we automatically matched the terminology comprised in the four intensional definitions with the MeSH terms. We list the terms that were automatically extracted in each definition:

- Definition 1: Kidney; Brain
- Definition 2: Tuberous Sclerosis; Intellectual Disability; Viscera; Brain; Seizures; Retina
- Definition 3: Humans; Tuberous Sclerosis Complex 1 Protein; Tuberous Sclerosis Complex 2 Protein; Tumor Suppressor Proteins; Kidney; Mutation; Brain; Neoplasms; Liver; Lung
- Definition 4: Intellectual Disability; Mutation; Brain; Seizures; Kidney

Also in this case, the comparison between the MeSH terms extracted from the definitions and the terms that compose at the semic level the corresponding formulations of different formulations of semic analysis, revealed that the term-matching process did not manage to extract all the terms that compose the semic analysis. As we previously mentioned, we expected the absence from the extracted MeSH terms of both the terms that represent the mesogeneric seme and the seme that specifies the subdomain in which the reference term is used. For what concerns the MeSH terms extracted in the first intensional definition, we observed that the term *Brain* correctly matches the seme */brain/* that is included in the corresponding sememe. On the contrary, the MeSH term *Kidney* was not included in the sememe in the form of semes. With reference to the MeSH terms that were extracted in the second definition, the term *Brain* also, in this circumstance, matches the seme */brain/* contained in the related sememe. Instead, the MeSH terms *Tuberous Sclerosis*, *Viscera*, *Seizures* and *Retina* are not present in the corresponding sememe in the form of semes. In particular, we noticed that the term *Intellectual Disability* was indicated as the MeSH term for the medical term *mental retardation*, with the latter being included in the second intensional definition. Subsequently, we examined the MeSH terms that were extracted in the third definition. Specifically, we observed that only the MeSH terms *Brain* and *Mutation* are included in the corresponding sememe in the form of semes, and that both match the respective semes, also from a lexical viewpoint. The analysis of the list of MeSH terms extracted in the fourth intensional definition resulted in the finding that the MeSH terms *Intellectual Disability*, *Mutation* and *Brain* are all included as semes in the related sememe. Particularly in this circumstance, the MeSH terms *Intellectual Disability* and *Mutation* both match the semes that, in the context of the sememe, constitute the subset of connotative semes of the reference term *tuberous sclerosis*.

4.2. Analysis of Polyreferentiality

In this section, we perform the semic analysis of two different terms: *capsule* and *resistance*. In this case, we aim to exemplify our approach to the analysis of the phenomenon of polyreferentiality. Moreover, we analyze the result of the automatic extraction of MeSH terms from the intensional definitions that constitute our corpus. The objective is to evaluate if an exact automatic extraction of the terms that constitute the semes of the reference terms is achieved.

4.2.1. Capsule

In the first place, we propose the semic analysis of the term *capsule*, which was performed by manually applying the proposed methodology. In this circumstance, the intensional definitions used as a corpus originated from the reformulation of the definitions contained in the Merriam–Webster Medical Dictionary, the TheFreeDictionary’s Medical Dictionary and Dictionary.com. The generation of each sememe, in particular, is conducted by considering more than one definition. An exception is constituted by the eighth sememe, for which a single definition has been found. In Appendix C we include the intensional definitions that we took as a reference. The listed sememes represent the final result of the analysis:

1. /medicine/ /anatomy/ /membrane/ /structure/ /part/ /organ/ /cartilaginous structure/ /fibrous structure/ /fatty structure/ /membranous structure/ /mem-

- branous anatomic structure/ /irregular/ /collagenous connective tissue/ /joint/ /capsule/ /envelope/ /tumor/ /benign/ /fibrous tissue layer/ /fibrous sheath/ /membranous sheath/ /fatty sheath/ /sac/ /kidney/ /fibrous tissue/ /fibromembranous/ /liver/ /fibrous/ /ensheathe/ /benign tumour/ /cyst/ /parasite/ /host defense/ /neoplasm/ /protective/ /coat/ /solid organ/ /spleen/ /membrane/ /crystalline lens/ /eye/ /stability/ /function/ /nephron/ /membranous sac/ /integument/ /bacterium/ /encapsulate/ /resistance/ /defense/ /host/
2. /medicine/ /anatomy/ /layer/ /lamina/ /white matter/ /cerebrum/ /stratum/
 3. /medicine/ /pharmacology/ /shell/ /gelatin/ /drug/ /vitamin/ /soluble/ /dose/ /medicine/ /gelatine/ /powder/ /liquid/ /structure/ /gelatinous case/
 4. /medicine/ /pharmacology/ /preparation/ /oral/ /medicinal/ /nutritional/ /shell/ /content/ /solid dosage/ /drug/ /soluble/ /gelatin/
 5. /medicine/ /microbiology/ /envelope/ /polysaccharide/ /viscous/ /gelatinous/ /microorganism/ /streptococcus pneumoniae/ /shell/ /bacterium/ /hyaline/ /polysaccharide/ /fungal wall/ /bacterial wall/ /fungal cell/ /bacterial cell/
 6. /medicine/ /botany/ /dry dehiscent fruit/ /carpel/ /fruit/ /angiosperm/ /dehiscent/
 7. /medicine/ /botany/ /structure/ /moss/ /plant/ /sporangium/ /bryophyte/ /layer/ /spore/ /organism/ /fern/ /alga/ /fungus/
 8. /medicine/ /anatomy/ /bone structure/ /vertebrate skull/

As can be observed, specialized terminology constitutes the denotative semes that compose the sememes, since the term *capsule* is specialized. With reference to the terms represented by the semes /structure/, /part/, /layer/ and /content/, they are not terms specific to medicine. However, these terms are included as entries in the Merriam–Webster Medical Dictionary. We added the mesogeneric seme /medicine/ that indicates the domain and the seme that represents the subdomain to each sememe. In particular, these sememes are composed by unique combinations of semes that differentiate one from another. As each unique combination of semes respectively corresponds to a different unique combination of characteristics, it follows that these sememes represent different concepts. Consequently, the terms that are represented at the semic level by these sememes are also different. As a matter of fact, these terms solely share the same lexical form, and they do not designate the same concept. Considering this, it is possible to identify polyreferentiality.

Subsequently, we present the analysis involving the matching of the MeSH terms extracted from the intensional definitions with the terms that we transposed into semes in the different sememes. Below, we list the MeSH terms that were extracted in the definitions:

- Definition 1: Cell membrane
- Definition 2: White Matter; Cerebrum
- Definition 3: Gelatin; Vitamins
- Definition 5: Streptococcus pneumoniae; Polysaccharides

The comparison between the extracted MeSH terms and the terms that compose at the semic level the corresponding semic analysis led to the observation that the automatic term matching did not manage to match and extract all the semes that compose the sememes. Also in this case, the mesogeneric seme and the seme that indicates the subdomain are absent, since they are not present in the formulated intensional definitions. Regarding the first definition, the term *Cell membrane* was extracted. However, the term does not match any of the semes included in the sememe. In this circumstance, therefore, an absence of extracted MeSH terms was observed. For what concerns the second intensional definition, the extracted MeSH terms *White Matter* and *Cerebrum* constitute two of the terms included as semes in the corresponding sememe. With regard to the MeSH terms related to the third definition, the term *Gelatin* matches the seme /gelatin/ comprised in the related sememe. The term *Vitamins*, instead, does not match the seme /vitamin/ from a lexical viewpoint. Particularly, it can be noticed that the number of MeSH terms extracted in the fourth definition amounts to zero. Instead, the MeSH term *Streptococcus pneumoniae*, extracted in the fifth definition, matches the seme /streptococcus pneumoniae/ included in the

corresponding sememe. However, the MeSH term *Polysaccharides* does not match the seme */polysaccharide/*.

4.2.2. Resistance

The same procedure was also adopted to conduct the semic analysis of the term *resistance*. We took as a reference the definitions of the term contained in the Merriam–Webster Medical Dictionary and in TheFreeDictionary’s Medical Dictionary. By reformulating the definitions contained in the resource, we obtained the intensional definitions of the term. Also in this circumstance, sememes are formulated by considering more than one definition. An exception is constituted by the second and the fifth sememes, for which a single definition has been found. Specifically, the intensional definitions are included in Appendix D. Subsequently, we generated the corresponding formulations of semic analysis, obtained by manually applying the proposed methodology:

1. /medicine/ /biology/ /power/ /capacity/ /resist/ /inherent/ /ability/ /organism/ /disease/ /toxic agent/ /infection/ /inherent/ /acquired/ /immunity/ /effect/ /antagonistic agent/ /toxin/ /drug/ /pathogenic microorganism/ /tissue/ /cell/ /destructive agent/ /destructive condition/ /chemical compound/ /disease agent/ /environmental stressor/ /host/ /pathogen/ /grow/ /presence/ /antibiotic/ /inherited/ /environmental factor/ /parasite/ /biocide/ /herbicide/ /insecticide/ /salinity/ /normal organism/ /noxious agent/ /environment/ /instance/
2. /medicine/ /biology/ /capacity/ /species/ /strain/ /microorganism/ /survive/ /exposure/ /toxic agent/ /drug/ /effective/ /genetic mutation/ /selection/ /accumulation/ /gene/ /protection/ /agent/ /overuse/ /selectively/ /protective gene/
3. /medicine/ /physics/ /body/ /passage/ /current/ /counteracting force/ /conductor/ /electricity/ /energy/ /agent/ /heat/ /voltage/
4. /medicine/ /physiology/ /impediment/ /flow/ /fluid/ /blood/ /respiratory gas/ /passage/ /blood flow/ /tracheobronchial tree/
5. /medicine/ /psychology/ /psychological defense mechanism/ /psychoanalysis patient/ /reject/ /therapeutic effort/ /analyst/
6. /medicine/ /psychiatry/ /conscious defense/ /unconscious defense/ /repressed/ /awareness/ /forgetfulness/ /evasion/ /mental block/ /anger/ /intellectualization/ /intensification/ /symptom/ /conscious psychologic defense/ /unconscious psychologic defense/ /recall/ /repressed unconscious thought/
7. /medicine/ /psychiatry/ /psychoanalysis/ /unconscious defense/ /repressed thought/ /consciousness/
8. /medicine/ /physics/ /force/ /active force/
9. /medicine/ /pathology/ /ability/ /erythrocyte/ /resist/ /hemolysis/ /preserve/ /shape/ /osmotic pressure/ /blood plasma/
10. /medicine/ /endocrinology/ /defective target tissue response/ /hormone/
11. /medicine/ /psychology/ /process/ /ego/ /conscious recall/ /anxiety/ /experience/
12. /medicine/ /biology/ /normal response/ /biologically active compound/ /hormone/
13. /medicine/ /oncology/ /failure/ /cancer/ /regress/ /RT/ /chemotherapy/
14. /medicine/ /physics/ /passive force/ /active force/

As can be noticed, since the medical term *resistance* is specialized, the terminology employed in the formulations of semic analysis is accordingly specialized. In particular, the sememes uniquely share the same mesogeneric seme */medicine/*. It follows that the sememes are respectively composed by a different unique combination of denotative semes, which correspond to a different unique combination of characteristics that make up the concept. Consequently, the phenomenon of polyreferentiality emerges. These sememes, therefore, represent at the semic level different terms, with each one designating a different concept. Indeed, as we previously mentioned, the condition that must be met for terms to differ from one another is diversity in at least one denotative seme.

Finally, we perform the analysis involving the matching of the MeSH terms extracted from the intensional definitions with the terms that we included as semes in the different sememes. For each definition, we list the matched MeSH terms:

- Definition 2: Mutation
- Definition 4: Gases
- Definition 5: Psychoanalysis; Defense Mechanisms

The comparison between the list of MeSH terms and the semes included in the corresponding sememes resulted in the ascertainment that the number of MeSH terms extracted in the first and in the third intensional definitions amounts to zero. For what concerns the other three intensional definitions of the reference term *resistance*, the MeSH terms that were extracted do not match the mesogeneric seme and the seme that refers to the subdomain. With reference to the MeSH terms extracted in the second definition, the term *Mutation* does not match the seme */genetic mutation/* contained in the corresponding sememe. The same circumstance applies to the MeSH term extracted in the fourth intensional definition, as the term *Gases* does not match the seme */respiratory gas/*. With regard to the MeSH terms related to the fifth definition, the term *Psychoanalysis* does not conceptually match the seme */psychoanalysis patient/* contained in the related sememe. Furthermore, the MeSH term *Defense Mechanism* does not match the seme */psychological defense mechanism/*.

5. Conclusions and Future Work

In this paper, we proposed a methodology for a systematic semic analysis of medical terminology that integrates, develops, and extends the methodology that was formerly presented by [Bonato et al. \(2021\)](#). The proposed procedure involves (1) the semi-automatic domain-specific corpus-based extraction of semes, (2) the application of the property of termhood to address the diaphasic and the diastratic variations of language, (3) the automatic lemmatization of semes, and (4) seme weighting to establish the order of semes in the sememe. One of the objectives of the study was to enhance the systematicity of semic analysis of medical terminology and to automate to a greater extent its performance. Our approach proved to be effective in increasing the systematicity of semic analysis of terminology that belongs to the medical domain. Indeed, the procedural criteria comprised in the methodology are purposefully intended to serve as a guideline, aiming to minimize subjectively driven phenomena in semic analysis of medical terms.

Additionally, the study aimed to propose an approach for analyzing polysemy and polyreferentiality within the framework of semic analysis of medical terminology. The case study demonstrated that the methodology constitutes a valuable approach for analyzing both polysemy and polyreferentiality in such a context. Specifically, we manually applied the first three procedural criteria of the methodology to perform the semic analysis of a set of four medical reference terms. The methodology enabled us to effectively identify the denotative semes and the connotative semes that are associated with the terms *Down syndrome* and *tuberous sclerosis*. Such identification made it possible to prove that multiple sememes can be associated with both reference terms: a sememe exclusively composed by denotative semes, and sememes composed by denotative semes and connotative semes. This circumstance led to the identification of the polysemic nature of the terms *Down syndrome* and *tuberous sclerosis*. Moreover, the identification of the connotative semes also made it possible to evaluate whether a negative or a positive connotation is attached to these terms. Furthermore, the application of the methodology enabled us to distinctively represent, at the semic level, the concepts that are linguistically designated by the terms *capsule* and *resistance*. Consequently, it was also possible to analyze the sets of different sememes that are respectively associated with these terms. In particular, we demonstrated that the sememes that are associated with the term *capsule* represent, at the semic level, concepts that are different from each another. Therefore, these sememes respectively designate different terms that uniquely share the same lexical form. The same finding also applied to the sememes of the term *resistance*. Hence, the analysis of the sememes that are

respectively associated with the terms *capsule* and *resistance* led to the assessment of the occurrence of the manifestation of polyreferentiality.

The proposed case study also concerned the application of an automatic procedure, which enabled us to match the medical terms included in the intensional definitions, taken as the basis for the performance of semic analysis, with the list of medical terms provided by the MeSH database. In particular, the objective was to evaluate if the automatic extraction of MeSH terms in the definitions can aid in identifying all the terms that constitute the semes of the reference terms. The analysis of the results revealed that the totality of the medical terms representing semes of the reference terms could not be automatically extracted and matched with the MeSH terms. Nevertheless, it can serve as a valuable resource to support the matching procedure for the creation of the sememes.

In future work, we propose to further develop the proposed methodology, and to conduct a case study that features the application of all the steps that are included in the methodology. Specifically, we intend to additionally perform some weighting to order the semes that compose the sememe by applying the *v-tech* value (Vezzani 2019). In this work, we will rely on a corpus of definitions extracted from lexicographic and terminological dictionaries. However, considering the limited dimension of our corpus, we also propose to rely on comprehensive corpus-based research to fully capture the dynamic and context-dependent nature of terms. Moreover, although we propose a methodology aimed at enhancing the systematicity and automaticity of semic analysis, we propose to, in future work, evaluate it by comparing the obtained results to a benchmark. Indeed, for what concerns semic analysis of medical terminology, such a benchmark is not currently available. Additionally, we will further research the parallel between the characteristics of a concept and the semantic primitives proposed by Wierzbicka and АННА (2021). Since semantic primitives are a set of universal, irreducible elements that form the building blocks of meaning in language, there might be a strong correlation between the primitives and the characteristics that form the unit of knowledge of a concept.

Finally, we propose to apply the proposed methodology in future studies focusing on the connotation that is attributed to terms within the medical domain.

Author Contributions: Conceptualization, V.B., F.V. and G.M.D.N.; methodology, V.B., F.V. and G.M.D.N.; writing—original draft preparation, V.B.; writing—review and editing, V.B., F.V. and G.M.D.N.; supervision, F.V. and G.M.D.N. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Informed Consent Statement: Not applicable.

Data Availability Statement: No new data were created or analyzed in this study. Data sharing is not applicable to this article.

Acknowledgments: This work is partially supported by the HEREDITARY Project, as part of the European Union's Horizon Europe research and innovation programme under grant agreement No GA 101137074, and it is part of the initiatives of the Center for Studies in Computational Terminology (CENTRICO) of the University of Padua and in the research directions of the Italian Common Language Resources and Technology Infrastructure CLARIN-IT.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Definition 1: Congenital condition characterized especially by developmental delays, usually mild to moderate impairment in cognitive functioning, short stature, relatively small head, upward slanting eyes usually with epicanthal folds, flattened nasal bridge, broad hands with short fingers, decreased muscle tone, and trisomy of the human chromosome numbered 21.⁹

Definition 2: Congenital condition characterized by physical malformations and some degree of mental retardation. It was formerly known as mongolism because of the fact

that the patient's facial characteristics resemble those of persons of the Mongolian race. Also called trisomy 21 syndrome, because the disorder is concerned with a defect in chromosome 21.¹⁰

Definition 3: Medical condition that some people are born with, caused by a fault with one chromosome, which results in learning disabilities.¹¹

Appendix B

Definition 1: Autosomal dominant multi-system disorder (OMIM:191100) characterized by hamartomas and developmental defects, which primarily affects the brain, kidneys, heart and skin.¹²

Definition 2: Congenital hereditary disease, transmitted as an autosomal dominant trait, characterized principally by the presence of hamartomas of the brain (tubers), retina (phakomas), and viscera, mental retardation, seizures, and adenoma sebaceum; often associated with other skin lesions.¹³

Definition 3: Disease that is rare that causes benign tumors to grow in the human brain, kidneys, heart, liver, eyes, lungs and skin; caused by a mutation in the genes for the tumor growth suppressor proteins hamartin and tuberlin.¹⁴

Definition 4: Genetic disorder of the skin and nervous system that is characterized by the formation of small benign tumors in various organs (such as the brain, kidney, eye, and heart). It is accompanied by variable symptoms, including seizures, developmental delay or intellectual disability, and skin lesions (as hypopigmented macules of the trunk and limbs or telangiectatic facial papules), and is inherited as an autosomal dominant trait or results from spontaneous mutation.¹⁵

Appendix C

Sememe 1

Definition 1: Membrane or structure that is saclike, which encloses a part or organ.¹⁶

Definition 2: Structure that is cartilaginous, fatty, fibrous or membranous, enveloping another structure, organ or part.¹⁷

Definition 3: Anatomic structure that is membranous; usually dense, irregular, collagenous connective tissue that envelops an organ, a joint or any other part resembling a capsule or envelope.¹⁸

Definition 4: Tissue layer that is fibrous, which envelops an organ or a tumor, especially if benign.¹⁹

Definition 5: A sheath that is fibrous, membranous or fatty, which encloses an organ or part, such as the sac surrounding the kidney or the fibrous tissues that surround a joint.²⁰

Definition 6: Covering of a fibromembranous nature of an organ, such as the kidney and the liver.²¹

Definition 7: Covering of a fibrous nature of a joint.²²

Definition 8: Material that is thick and fibrous, which ensheathes benign tumours, cysts or parasites that have been "walled off" as a host defense.²³

Definition 9: Tissue layer that is fibrous, which envelops an organ, joint or neoplasm.²⁴

Definition 10: Outer covering of any kind, such as the tough, protective outer coat of solid organs, including the kidneys, liver and spleen; the delicate outer membrane of the internal crystalline lens of the eye; or capsules of the joints that contribute to their stability and function.²⁵

Definition 11: Anatomic structure that is membranous, and is usually dense, irregular, collagenous connective tissue that envelops a body part.²⁶

Definition 12: Containing structure, with a strong outer covering and is, found in many different groups, such as the blind-ending part of the kidney nephron.²⁷

Definition 13: Sac that is membranous or integument.²⁸

Definition 14: Containing structure that, with a strong outer covering, and is found in many different groups, such as the outer coat of some bacteria (referred to as encapsulated), which enhances resistance to the defenses of the host.²⁹

Sememe 2

Definition 1: Either of two layers or laminae of white matter in the cerebrum.³⁰

Definition 2: Either of two strata of white matter in the cerebrum.³¹

Sememe 3

Definition 1: Shell, usually of gelatin; for packaging something, such as a drug or vitamins.³²

Definition 2: Container that is small and soluble, and usually made of gelatin, which encloses a dose of an oral medicine or a vitamin.³³

Definition 3: Containers that are soluble and made of gelatine, used for drugs in powder or liquid form.³⁴

Definition 4: Structure that is enclosing, a soluble container enclosing a dose of medicine.³⁵

Definition 5: Case that is gelatinous, which encloses a dose of medicine.³⁶

Sememe 4

Definition 1: Preparation for oral use that is usually medicinal or nutritional, consisting of the shell and its contents.³⁷

Definition 2: Dosage form of a solid nature of a drug, in which the drug is enclosed in a hard or soft soluble container or shell of an appropriate gelatin.³⁸

Definition 3: Dosage form of a solid nature, in which a drug is enclosed in either a hard or soft shell of soluble material.³⁹

Definition 4: Dosage form of a solid nature, in which a drug is enclosed in a hard or soft soluble container or "shell" of a suitable form of gelatin.⁴⁰

Sememe 5

Definition 1: Envelope that is polysaccharide, often viscous or gelatinous, which surrounds certain microscopic organisms (such as the pneumococcus).⁴¹

Definition 2: Outer shell that is polysaccharide, enveloping certain bacteria.⁴²

Definition 3: Polysaccharide that is hyaline, coating around a fungal or bacterial wall of a cell.⁴³

Definition 4: Polysaccharide that is hyaline, coating around a fungal or bacterial cell.⁴⁴

Sememe 6

Definition 1: Fruit that is dehiscent and dry, which develops from two or more united carpels.⁴⁵

Definition 2: Type of fruit of angiosperms that splits open when dry (dehiscent).⁴⁶

Sememe 7

Definition 1: Structure that is thin-walled and spore-containing of mosses and related plants.⁴⁷

Definition 2: Containing structure, with a strong outer covering found in many different groups, such as the sporangium of bryophytes (e.g., mosses of a hard outer layer inside which are developing spores).⁴⁸

Definition 3: Sporangium of various spore-producing organisms, such as ferns, mosses, algae, and fungi.⁴⁹

Sememe 8

Definition 1: Bony structure that is spherical, found in some vertebrate skulls.⁵⁰

Appendix D*Sememe 1*

Definition 1: Power or capacity to resist, and especially the inherent ability of an organism to resist harmful influences (such as disease, toxic agents, or infection).⁵¹

Definition 2: Ability, which is natural or acquired, of an organism to maintain its immunity to, or oppose the effects of, an antagonistic agent (for example, a toxin, drug, or pathogenic microorganism).⁵²

Definition 3: Ability of an organism, tissue or cell to withstand a destructive agent or condition, such as a chemical compound, a disease agent or an environmental stressor.⁵³

Definition 4: Ability of a host to resist a pathogen; able to grow in the presence of a particular antibiotic.⁵⁴

Definition 5: Characteristic of any kind that is inherited by an organism, which lessens the effect of an adverse environmental factor, such as a pathogen or parasite, a biocide (e.g., herbicide, insecticide, antibiotic) or a natural climatic extreme, such as drought or high salinity.⁵⁵

Definition 6: Ability, of a natural kind, of a normal organism to remain unaffected by noxious agents in its environment.⁵⁶

Definition 7: Act, an instance of resisting or the capacity to resist.⁵⁷

Definition 8: Ability, which is natural or acquired, of an organism to maintain its immunity to, or resist the effects of, an antagonistic agent, such as a pathogenic microorganism, a toxin or a drug.⁵⁸

Definition 9: Ability of an organism to maintain its immunity to, or oppose the effects of, an antagonistic agent.⁵⁹

Sememe 2

Definition 1: Capacity of a species or strain of microorganism to survive exposure to a toxic agent (such as a drug) formerly effective against it, due to genetic mutation, selection for and accumulation of genes conferring protection from the agent especially as a result of overuse of the agent which selectively destroys individual microorganisms lacking the protective genes.⁶⁰

Sememe 3

Definition 1: Opposition offered by a body to the passage through it of a steady electric current.⁶¹

Definition 2: Opposition or counteracting force, such as opposition of a conductor to the passage of electricity or another energy or substance.⁶²

Definition 3: Opposition in a conductor to the passage of a current of electricity, whereby there is a loss of energy and a production of heat; specifically, the potential difference in volts across the conductor per ampere of current flow.⁶³

Sememe 4

Definition 1: Opposition or impediment to the flow of a fluid, such as blood or respiratory gases, through one or more passages.⁶⁴

Definition 2: Opposition to flow of a fluid through one or more passageways—for example, blood flow and respiratory gases in the tracheobronchial tree.⁶⁵

Definition 3: Opposition to the flow of a fluid through one or more passageways.⁶⁶

Sememe 5

Definition 1: Psychological defense mechanism wherein a psychoanalysis patient rejects, denies or otherwise opposes therapeutic efforts by the analyst.⁶⁷

Sememe 6

Definition 1: Defenses, which are conscious or unconscious, against change and intended to prevent repressed material from coming into awareness. It can take such forms as forgetfulness, evasions, embarrassment, mental blocks, denial, anger, superficial talk, intellectualization or intensification of symptoms.⁶⁸

Definition 2: Defense that is psychologic, conscious or unconscious, against recall of repressed unconscious thoughts.⁶⁹

Sememe 7

Definition 1: Defense in psychoanalysis, which is adopted by a person in an unconscious way, to protect against bringing repressed thoughts to consciousness.⁷⁰

Definition 2: Defense, of an unconscious nature, by a person against bringing repressed thoughts to consciousness.⁷¹

Sememe 8

Definition 1: Force exerted in opposition to an active force.⁷²

Sememe 9

Definition 1: Ability of red blood cells to resist hemolysis and to preserve their shape under varying degrees of osmotic pressure in the blood plasma.⁷³

Sememe 10

Definition 1: Target tissue response that is defective to a hormone, in endocrinology.⁷⁴

Sememe 11

Definition 1: Process in which the ego opposes the conscious recall of anxiety-producing experiences.⁷⁵

Sememe 12

Definition 1: Lack of normal response to a biologically active compound, such as a hormone.⁷⁶

Sememe 13

Definition 1: Failure of a cancer to regress after RT or chemotherapy.⁷⁷

Sememe 14

Definition 1: Force that is passive, exerted in opposition to another active force.⁷⁸

Notes

¹ In the original formulation of the semic analysis performed by Stan (2014, p. 71), the seme /+shape/ is included in the sememe. In particular, Stan (2014, pp. 71–72) affirms that “the physical defects can affect parts of the body (i) by the manifestation of a disorder / dysfunction, (ii) by the appearance or disappearance of certain components or (iii) by the modification of the aspect of a certain component. These three characteristics can be abstracted by means of the semes (i) /+disorder/ /-ability/, (ii) /+component(s)/ /+added/ /+ reduced/ and (iii) /+shape/”. In the study by the author, therefore, the graphic signs “+” or “–” accordingly express these variables. With the purpose of normalising the metalinguistic representation of semic analysis in our contribution, however, we omitted the graphic sign “+” that was originally included by the author within the seme /+shape/.

² <https://www.merriam-webster.com/medical> (accessed on 21 March 2024).

³ <https://medical-dictionary.thefreedictionary.com/> (accessed on 21 March 2024).

⁴ <https://cran.r-project.org/web/packages/tidytext/> (accessed on 21 March 2024).

⁵ <https://www.ncbi.nlm.nih.gov/mesh> (accessed on 21 March 2024).

⁶ <https://www.nlm.nih.gov/oet/ed/mesh/meshondemand.html> (accessed on 21 March 2024).

⁷ <https://www.oxfordlearnersdictionaries.com/> (accessed on 21 March 2024).

⁸ <https://www.yourdictionary.com/> (accessed on 21 March 2024).

⁹ Merriam-Webster Medical Dictionary: Down syndrome. <https://www.merriam-webster.com/dictionary/Down%20syndrome#medicalDictionary> (accessed on 21 March 2024).

¹⁰ TheFreeDictionary’s Medical Dictionary: Down syndrome. *Miller-Keane Encyclopedia and Dictionary of Medicine, Nursing, and Allied Health, Seventh Edition*, Saunders. <https://medical-dictionary.thefreedictionary.com/Down+syndrome> (accessed on 21 March 2024).

¹¹ Oxford Learner’s Dictionary: Down syndrome. <https://www.oxfordlearnersdictionaries.com/definition/english/down-syndrome?q=Down+syndrome> (accessed on 21 March 2024).

¹² TheFreeDictionary’s Medical Dictionary: tuberous sclerosis. *Segen’s Medical Dictionary*, Farlex Inc. <https://medical-dictionary.thefreedictionary.com/tuberous+sclerosis> (accessed on 21 March 2024).

¹³ TheFreeDictionary’s Medical Dictionary: tuberous sclerosis. *Miller-Keane Encyclopedia and Dictionary of Medicine, Nursing, and Allied Health, Seventh Edition*, Saunders. <https://medical-dictionary.thefreedictionary.com/tuberous+sclerosis> (accessed on 21 March 2024).

¹⁴ YourDictionary: tuberous sclerosis. <https://www.yourdictionary.com/tuberous-sclerosis> (accessed on 21 March 2024).

¹⁵ Merriam-Webster Medical Dictionary: tuberous sclerosis. <https://www.merriam-webster.com/medical/tuberous%20sclerosis> (accessed on 21 March 2024).

¹⁶ Merriam-Webster Medical Dictionary: capsule. <https://www.merriam-webster.com/dictionary/capsule#medicalDictionary> (accessed on 21 March 2024).

¹⁷ TheFreeDictionary’s Medical Dictionary: capsule. *Miller-Keane Encyclopedia and Dictionary of Medicine, Nursing, and Allied Health, Seventh Edition*, Saunders. <https://medical-dictionary.thefreedictionary.com/capsule> (accessed on 21 March 2024).

¹⁸ TheFreeDictionary’s Medical Dictionary: capsule. *Farlex Partner Medical Dictionary*. Farlex. <https://medical-dictionary.thefreedictionary.com/capsule> (accessed on 21 March 2024).

¹⁹ see footnote 18.

²⁰ TheFreeDictionary’s Medical Dictionary: capsule. *The American Heritage Medical Dictionary*. Houghton Mifflin Company. <https://medical-dictionary.thefreedictionary.com/capsule> (accessed on 21 March 2024).

²¹ TheFreeDictionary’s Medical Dictionary: capsule. *Segen’s Medical Dictionary*. Farlex. <https://medical-dictionary.thefreedictionary.com/capsule> (accessed on 21 March 2024).

²² see footnote 21.

²³ see footnote 21.

- 24 TheFreeDictionary's Medical Dictionary: capsule. *Medical Dictionary for the Health Professions and Nursing*. Farlex. <https://medical-dictionary.thefreedictionary.com/capsule> (accessed on 21 March 2024).
- 25 TheFreeDictionary's Medical Dictionary: capsule. *Collins Dictionary of Medicine*. Robert M. Youngson. <https://medical-dictionary.thefreedictionary.com/capsule> (accessed on 21 March 2024).
- 26 TheFreeDictionary's Medical Dictionary: capsule. *Medical Dictionary for the Dental Professions*. Farlex. <https://medical-dictionary.thefreedictionary.com/capsule> (accessed on 21 March 2024).
- 27 TheFreeDictionary's Medical Dictionary: capsule. *Collins Dictionary of Biology, 3rd edition*. W. G. Hale, V. A. Saunders, J. P. Margham. <https://medical-dictionary.thefreedictionary.com/capsule> (accessed on 21 March 2024).
- 28 Dictionary.com: capsule. <https://www.dictionary.com/browse/capsule> (accessed on 21 March 2024).
- 29 see footnote 27.
- 30 see footnote 16.
- 31 see footnote 28.
- 32 see footnote 16.
- 33 see footnote 20.
- 34 see footnote 25.
- 35 see footnote 17.
- 36 see footnote 28.
- 37 see footnote 16.
- 38 TheFreeDictionary's Medical Dictionary: capsule. *McGraw-Hill Concise Dictionary of Modern Medicine*. The McGraw-Hill Companies, Inc. <https://medical-dictionary.thefreedictionary.com/capsule> (accessed on 21 March 2024).
- 39 see footnote 24.
- 40 see footnote 26.
- 41 see footnote 16.
- 42 see footnote 20.
- 43 see footnote 24.
- 44 see footnote 18.
- 45 see footnote 20.
- 46 see footnote 27.
- 47 see footnote 20.
- 48 see footnote 27.
- 49 see footnote 28.
- 50 see footnote 27.
- 51 Merriam-Webster Medical Dictionary: resistance. <https://www.merriam-webster.com/dictionary/resistance#medicalDictionary> (accessed on 21 March 2024).
- 52 TheFreeDictionary's Medical Dictionary: resistance. *Farlex Partner Medical Dictionary*. Farlex. <https://medical-dictionary.thefreedictionary.com/resistance> (accessed on 21 March 2024).
- 53 TheFreeDictionary's Medical Dictionary: resistance. *The American Heritage Medical Dictionary*. Houghton Mifflin Company. <https://medical-dictionary.thefreedictionary.com/resistance> (accessed on 21 March 2024).
- 54 TheFreeDictionary's Medical Dictionary: resistance. *McGraw-Hill Concise Dictionary of Modern Medicine*. The McGraw-Hill Companies, Inc. <https://medical-dictionary.thefreedictionary.com/resistance> (accessed on 21 March 2024).
- 55 TheFreeDictionary's Medical Dictionary: resistance. *Collins Dictionary of Biology, 3rd edition*. W. G. Hale, V. A. Saunders, J. P. Margham. <https://medical-dictionary.thefreedictionary.com/resistance> (accessed on 21 March 2024).
- 56 TheFreeDictionary's Medical Dictionary: resistance. *Miller-Keane Encyclopedia and Dictionary of Medicine, Nursing, and Allied Health, Seventh Edition*. Saunders. <https://medical-dictionary.thefreedictionary.com/resistance> (accessed on 21 March 2024).
- 57 see footnote 53.
- 58 TheFreeDictionary's Medical Dictionary: resistance. *Medical Dictionary for the Health Professions and Nursing*. Farlex. <https://medical-dictionary.thefreedictionary.com/resistance> (accessed on 21 March 2024).
- 59 TheFreeDictionary's Medical Dictionary: resistance. *Medical Dictionary for the Dental Professions*. Farlex. <https://medical-dictionary.thefreedictionary.com/resistance> (accessed on 21 March 2024).
- 60 see footnote 51.
- 61 see footnote 51.
- 62 see footnote 56.
- 63 see footnote 52.
- 64 see footnote 51.
- 65 see footnote 52.
- 66 see footnote 58.
- 67 see footnote 51.

- 68 see footnote 56.
 69 see footnote 54.
 70 see footnote 52.
 71 see footnote 58.
 72 see footnote 52.
 73 see footnote 52.
 74 see footnote 52.
 75 see footnote 53.
 76 see footnote 53.
 77 see footnote 54.
 78 see footnote 58.

References

- Baider, Fabienne H. 2009. Lexical change, discourse practices and the French press. Plus ça change, plus c'est la même chose? In *Language Variation—European Perspectives II: Selected Papers from the 4th International Conference on Language Variation in Europe (ICLaVE 4), Nicosia, June 2007*. Edited by Stavroula Tsiplakou, Marilena Karyolemou and Pavlos Pavlou. Amsterdam and Philadelphia: John Benjamins Publishing Company, pp. 27–46. [\[CrossRef\]](#)
- Baider, Fabienne. 2014. Analysing Journalistic Discourse and Finding Opinions Semi-Automatically? A Case Study of the 2007 and 2012 French Presidential Campaigns. *Journal of Data Mining and Digital Humanities*. [\[CrossRef\]](#)
- Berruto, Gaetano. 2010. Identifying dimensions of linguistic variation in a language space. In *Language and Space; An International Handbook of Linguistic Variation. Vol. 1: Theories and Methods*. Edited by Peter Auer and Jürgen Erich Schmidt. Berlin: De Gruyter Mouton, pp. 226–41. [\[CrossRef\]](#)
- Bertels, Ann. 2011. The dynamics of terms and meaning in the domain of machining terminology. *Terminology. International Journal of Theoretical and Applied Issues in Specialized Communication* 17: 94–112. [\[CrossRef\]](#)
- Bianca, Mariano L., and Paolo Piccari. 2007. Inherent Logic: Isotopic and Inherent Bonds in Argumentation. *Anthropology & Philosophy* 8: 9–31.
- Bloomfield, Leonard. 1933. *Language*. London: George Allen & Unwin.
- Bonato, Vanessa, Giorgio Maria Di Nunzio, and Federica Vezzani. 2021. Preliminary Considerations on a Systematic Approach to Semic Analysis: The Case Study of Medical Terminology. *Umanistica Digitale* 5: 211–34. [\[CrossRef\]](#)
- Buysschaert, Joost. 2021. Medical terminology and discourse. In *The Routledge Handbook of Translation and Health*. Edited by Şebnem Susam-Saraeva and Eva Spišáková. London: Routledge, pp. 65–79. [\[CrossRef\]](#)
- Carvalho, Sara, Rute Costa, and Christophe Roche. 2016. LESS Can Indeed Be More: Linguistic and Conceptual Challenges in the Age of Interoperability. In *Proceedings of the 12th International Conference on Terminology and Knowledge Engineering*. Edited by Hanne Erdman Thomsen, Antonio Pareja-Lora and Bodil Nistrup Madsen. Copenhagen: Copenhagen Business School, pp. 157–67.
- Cluley, Victoria, Alison Pilnick, and Rachel Fyson. 2022. Talking about learning disability: Discursive acts in managing an ideological dilemma. In *SSM—Qualitative Research in Health*. Amsterdam: Elsevier, vol. 2. [\[CrossRef\]](#)
- Condamines, Anne, and Josette Rebeyrolle. 1997. Point de vue en langue spécialisée. *Meta: Journal des Traducteurs/Meta: Translators' Journal* 42: 174–84. [\[CrossRef\]](#)
- Conso, Danièle. 2020. L'apport de l'analyse sémique dans un cas de polysémie complexe (latin forma). *Revue de Linguistique Latine du Centre Alfred Ernout (De Lingua Latina)* 20: 95–117.
- Corbeil, Jean-Claude. 1988. Les terminologies devant Babel. In *Actes du Colloque Terminologie et Technologies Nouvelles*. La Défense, du 9 au 11 décembre 1985. Québec: Office de la Langue Française, pp. 49–62.
- Cortelazzo, Michele. 1994. *Lingue Speciali. La Dimensione Verticale*, 2nd ed. Padova: Unipress.
- Costa, Rute. 2013. Terminology and Specialised Lexicography: Two Complementary Domains. *Lexicographica* 29: 29–42. [\[CrossRef\]](#)
- Cruse, D. Alan. 2001. Microsenses, default specificity and the semantics-pragmatics boundary. *Axiomathes* 12: 35–54. [\[CrossRef\]](#)
- Cusimano, Christophe. 2007. Polysémie et noms de sentiments: Essai de sémantique générale. Thèse de doctorat, Université de Metz, Metz, France.
- Dendien, Jacques, and Jean-Marie Pierrel. 2003. Le Trésor de la Langue Française Informatisé: Un exemple d'informatisation d'un dictionnaire de langue de référence. *Traitement Automatique des Langues* 44: 11–37.
- Di Nunzio, Giorgio Maria, and Federica Vezzani. 2018. Using R Markdown for Replicable Experiments in Evidence Based Medicine. In *Experimental IR Meets Multilinguality, Multimodality, and Interaction. 9th International Conference of the CLEF Association, CLEF 2018, Avignon, France, September 10–14, 2018, Proceedings*. Edited by Patrice Bellot, Chiraz Trabelsi, Josiane Mothe, Fionn Murtagh, Jian Yun Nie, Laure Soulier, Eric SanJuan, Linda Cappellato and Nicola Ferro. Cham: Springer, pp. 28–39. [\[CrossRef\]](#)
- Di Nunzio, Giorgio Maria, and Federica Vezzani. 2021. On the Reusability of Terminological Data. In *AIUCD 2021—DH per la società: E-guaglianza, partecipazione, diritti e valori nell'era digitale. Raccolta degli abstract estesi della 10ª conferenza nazionale*. Edited by Federico Boschetti, Angelo Mario Del Grosso and Enrica Salvatori. Pisa: Associazione per l'Informatica Umanistica e la Cultura Digitale, pp. 183–86.

- Dong, Zhendong, and Qiang Dong. 2003. HowNet—A hybrid language and knowledge resource. In *International Conference on Natural Language Processing and Knowledge Engineering, 2003. Proceedings. 2003*. Edited by Chengqing Zong. Piscataway: IEEE, pp. 820–24. [\[CrossRef\]](#)
- Dong, Zhendong, and Qiang Dong. 2006. *HowNet and the Computation of Meaning*. Singapore: World Scientific Publishing Co. Pte. Ltd. [\[CrossRef\]](#)
- Du, Jiaju, Fanchao Qi, Maosong Sun, and Zhiyuan Liu. 2020. Lexical Sememe Prediction using Dictionary Definitions by Capturing Local Semantic Correspondence. *arXiv* arXiv:2001.05954.
- Dury, Pascaline. 2013. Que montre l'étude de la variation d'une terminologie dans le temps. Quelques pistes de réflexion appliquées au domaine médical. *Debate Terminológico* 9: 2–10.
- Džuganová, Božena. 2017. Some Aspects of Medical English Terminology. *Foreign Language for Special Purposes* 5: 93–105.
- Filipec, Josef. 1994. Lexicology and Lexicography: Development and State of the Research. In *The Prague School of Structural and Functional Linguistics*. Edited by Philip A. Luelsdorff. Amsterdam and Philadelphia: John Benjamins Publishing Company, pp. 163–83.
- Frassi, Paolo. 2022. La représentation de la polysémie et des termes complexes de type *locution faible* dans une base de données terminologique. *Travail et son entourage dans le domaine du commerce international. Terminology. International Journal of Theoretical and Applied Issues in Specialized Communication* 28: 103–28. [\[CrossRef\]](#)
- Freixa, Judit, and Sabela Fernández-Silva. 2017. Terminological Variation and the Unsaturation of Concepts. In *Multiple Perspectives on Terminological Variation*. Edited by Patrick Drouin, Aline Francœur, John Humbley and Aurélie Picton. Amsterdam and Philadelphia: John Benjamins Publishing Company, pp. 155–80. [\[CrossRef\]](#)
- Geeraerts, Dirk. 1993. Vagueness's Puzzles, Polysemy's Vagaries. *Cognitive Linguistics* 4: 223–72. [\[CrossRef\]](#)
- Gotti, Maurizio. 2008. *Investigating Specialized Discourse*, 2nd revised ed. Bern: Peter Lang.
- Grzesitchak, Mick, Evelyne Jacquy, and Mathieu Valette. 2007. Systèmes complexes et analyse textuelle: Traits sémantiques et recherche d'isotopies. In *Actes du colloque ARCo'07: Cognition, Complexité, Collectif*. Nancy-Loria: Acta-Cognitica, pp. 227–35.
- Hébert, Louis. 2020. *An Introduction to Applied Semiotics: Tools for Text and Image Analysis*. Abingdon and New York: Routledge.
- ISO 1087: 2019. 2019, Terminology Work and Terminology Science—Vocabulary. Geneva: International Organization for Standardization.
- Jarvik, Gail P., and James P. Evans. 2017. Mastering genomic terminology. *Genetics in Medicine* 19: 491–92. [\[CrossRef\]](#) [\[PubMed\]](#)
- Kageura, Kyo, and Bin Umino. 1996. Methods of Automatic Term Recognition: A Review. *Terminology. International Journal of Theoretical and Applied Issues in Specialized Communication* 3: 259–89. [\[CrossRef\]](#)
- L'Homme, Marie-Claude. 2020a. *Lexical Semantics for Terminology: An Introduction*. Amsterdam and Philadelphia: John Benjamins Publishing Company. [\[CrossRef\]](#)
- L'Homme, Marie-Claude. 2020b. Revisiting Polysemy in Terminology. In *Proceedings of XIX EURALEX Congress: Lexicography for Inclusion*. Edited by Zoe Gavriilidou, Maria Mitsiaki and Asimakis Fliatouras. Greece: Democritus University of Thrace, pp. 415–24.
- L'Homme, Marie-Claude. 2023. Managing polysemy in terminological resources. In *Terminology. International Journal of Theoretical and Applied Issues in Specialized Communication*. Amsterdam: John Benjamins Publishing Company. [\[CrossRef\]](#)
- Laurenco, Robert. 2017. *Foundations for Clinical Neurology*. New York: Oxford University Press.
- León-Araúz, Pilar, and Antonio San Martín. 2011. Distinguishing Polysemy from Contextual Variation in Terminological definitions. In *Actas del X Congreso de la Asociación Europea de Lenguas para Fines Específicos: La investigación y la enseñanza aplicadas a las lenguas de especialidad y a la tecnología*. Edited by María Luisa Carrió, Josefa Contreras, Françoise Olmo, Hanna Skorczynska, Inmaculada Tamarit and Debra Westall. Valencia: Universitat Politècnica de València, pp. 173–86.
- León-Araúz, Pilar, Arianne Reimerink, and Alejandro García Aragón. 2013. Dynamism and context in specialized knowledge. *Terminology. International Journal of Theoretical and Applied Issues in Specialized Communication* 19: 31–61. [\[CrossRef\]](#)
- Li, Wei, Xuancheng Ren, Damai Dai, Yunfang Wu, Houfeng Wang, and Xu Sun. 2018. Sememe Prediction: Learning Semantic Knowledge from Unstructured Textual Wiki Descriptions. *arXiv* arXiv:1808.05437.
- Liu, Haibin, Tom Christiansen, William A. Baumgartner, Jr., and Karin Verspoor. 2012. BioLemmatizer: A lemmatization tool for morphological processing of biomedical text. *Journal of Biomedical Semantics* 3: 3. [\[CrossRef\]](#) [\[PubMed\]](#)
- Liu, Zhiyuan, Yankai Lin, and Maosong Sun. 2020. Sememe Knowledge Representation. In *Representation Learning for Natural Language Processing*. Singapore: Springer Singapore, pp. 125–61. [\[CrossRef\]](#)
- Miller, George A. 1995. WordNet: A lexical database for English. *Communications of the ACM* 38: 39–41. [\[CrossRef\]](#)
- Navigli, Roberto, and Simone Paolo Ponzetto. 2012. BabelNet: The automatic construction, evaluation and application of a wide-coverage multilingual semantic network. *Artificial Intelligence* 193: 217–50. [\[CrossRef\]](#)
- Niu, Yilin, Ruobing Xie, Zhiyuan Liu, and Maosong Sun. 2017. Improved Word Representation Learning with Sememes. In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. Edited by Regina Barzilay and Min-Yen Kan. Association for Computational Linguistics. Stroudsburg: Association for Computational Linguistics, pp. 2049–58. [\[CrossRef\]](#)
- Peláez, Jesús, and Juan Mateos. 2018. *New Testament Lexicography: Introduction—Theory—Method*. Edited by David S. du Toit. Berlin/Boston: De Gruyter. [\[CrossRef\]](#)
- Pethő, Gergely. 2001. What is polysemy?—A Survey of Current Research and Results. In *Pragmatics and the Flexibility of Word Meaning*. Edited by Károly Bibok and Eniko T. Németh. Amsterdam: Elsevier, pp. 175–224. [\[CrossRef\]](#)

- Picoche, Jacqueline. 1986. *Structures sémantiques du lexique français*. Paris: Fernand Nathan.
- Pottier, Bernard. 1992. The Components of Communication. In *Ideas, Words and Things. French Writings in Semiology*. Edited by Harjeet Singh Gill and Bernard Pottier. New Delhi: Orient Longman, pp. 113–35.
- Prié, Yannick. 1995. *Contribution à une clarification des rapports entre Sémantique Interprétative et Informatique*. Rennes: DEA Informatique—Université de Rennes I.
- Qi, Fanchao, Chuancheng Lv, Zhiyuan Liu, Xiaojun Meng, Maosong Sun, and Hai-Tao Zheng. 2022. Sememe Prediction for BabelNet Synsets using Multilingual and Multimodal Information. In *Findings of the Association for Computational Linguistics: ACL 2022*. Edited by Smaranda Muresan, Preslav Nakov and Aline Villavicencio. Association for Computational Linguistics: pp. 158–68. [CrossRef]
- Qi, Fanchao, Yangyi Chen, Fengyu Wang, Zhiyuan Liu, Xiao Chen, and Maosong Sun. 2021. Automatic Construction of Sememe Knowledge Bases via Dictionaries. In *Findings of the Association for Computational Linguistics: ACL-IJCNLP 2021*. Edited by Chengqing Zong, Fei Xia, Wenjie Li and Roberto Navigli. Stroudsburg: Association for Computational Linguistics, pp. 4673–86. [CrossRef]
- Qin, Yujia, Zhiyuan Liu, Yankai Lin, and Maosong Sun. 2023. Sememe-Based Lexical Knowledge Representation Learning. In *Representation Learning for Natural Language Processing*. Edited by Zhiyuan Liu, Yankai Lin and Maosong Sun. Singapore: Springer Singapore, pp. 351–400. [CrossRef]
- Rastier, François. 1987. *Sémantique interprétative*. Paris: Presses Universitaires de France.
- Rastier, François. 2005. La microsémantique. *Texto!* vol. 10. Available online: http://www.revue-texto.net/Inedits/Rastier/Rastier_Microsemantique.html (accessed on 17 March 2024).
- Rastier, François. 2014. La polysémie existe-t-elle? Quelques doutes constructifs. *Études romanes de Brno* 35: 23–39.
- Rastier, François. 2015. *Sémantique et recherches cognitives*. Paris: Presses Universitaires de France.
- Reutenauer, Coralie, Evelyne Jacquey, Michelle Lecolle, and Mathieu Valette. 2010. Sémème au microscope: Genèse et variation sémiques d'une unité lexicale. In *Statistical Analysis of Textual Data. Proceedings of 10th International Conference Journées d'Analyse statistique des Données Textuelles. 9–11 June 2010—Sapienza University of Rome*. Edited by Sergio Bolasco, Isabella Chiari and Luca Giuliano. Milano: LED, pp. 467–78.
- Reutenauer, Coralie, Mathieu Valette, and Evelyne Jacquey. 2009. De l'annotation sémique globale à l'interprétation locale: Environnement et image sémiques d' «économie réelle» dans un corpus sur la crise financière. *Cognitica—Actes du colloque de l'Association pour la Recherche Cognitive ARCo'09: Interprétation et problématiques du sens*; Rouen, France, pp. 29–39. Available online: https://hal.science/hal-00441166/file/Soumission_ARCo_corrige_13_oct_21h22.pdf (accessed on 17 March 2024).
- Reutenauer, Coralie. 2012. *Vers un traitement automatique de la néosémie: Approche textuelle et statistique*. Thèse de doctorat, Université de Lorraine, Metz, France.
- Roche, Christophe. 2012. Should Terminology Principles be re-examined? In *Proceedings of the 10th Terminology and Knowledge Engineering Conference: New Frontiers in the Constructive Symbiosis of Terminology and Knowledge Engineering (TKE 2012)*. Edited by Guadalupe Aguado de Cea, Mari Carmen Suárez-Figueroa, Raúl García-Castro and Elena Montiel-Ponsoda. Madrid: Universidad Politécnica de Madrid, pp. 17–32.
- Rouleau, Maurice. 2003. La terminologie médicale et ses problèmes. *Panace@* 4: 143–52.
- Santos, Claudia, and Rute Costa. 2015. Domain Specificity: Semasiological and Onomasiological Knowledge Representation. In *Handbook of Terminology—Volume 1*. Edited by Hendrik J. Kockaert and Frieda Steurs. Amsterdam and Philadelphia: John Benjamins Publishing Company, pp. 153–79. [CrossRef]
- Snipstad, Øyvind Ibrahim Marøy. 2022. Concerns regarding the use of the vulnerability concept in research on people with intellectual disability. *British Journal of Learning Disabilities* 50: 107–14. [CrossRef]
- Staicu, Simona Nicoleta. 2021. Semantic Structure of Some Causative Verbs in Medical Terminology. A Semic Analysis. In *Reading Multiculturalism. Human and Social Perspectives*. Edited by Iulian Boldea, Cornel Sigmirean and Dumitru-Mircea Buda. Tirgu Mures: Arhipelag XXI Press, pp. 38–45.
- Stan, Alexandra. 2014. A Semic Analysis of Terms Denoting Physical Defects in Romanian. In *10th Conference on British and American Studies—Crossing Boundaries: Approaches to the Contemporary Multicultural Discourse*. Edited by Marinela Burada and Oana Tatu. Newcastle upon Tyne: Cambridge Scholars Publishing, pp. 65–85.
- Sterner, Beckett. 2022. Explaining ambiguity in scientific language. *Synthese* 200: 1–27. [CrossRef]
- Thomas, Jean-François. 2017. Syntaxe et polysémie verbale: L'exemple de préverbés latins en de. *Pallas* 103: 105–13. [CrossRef]
- Thomas, Jean-François. 2023. Introduction aux études sémantiques dans le cadre de l'analyse sémique. Questions de méthode. *Revue de Linguistique Latine du Centre Alfred Ernout (De Lingua Latina)* 23: 1–6. Available online: <https://univ-montpellier3-paul-valery.hal.science/hal-04027353/> (accessed on 17 March 2024).
- Touratier, Christian. 2000. *La Sémantique*. Paris: Armand Colin.
- Veale, Tony. 2005. Analogy generation with HowNet. In *IJCAI'05: Proceedings of the 19th International Joint Conference on Artificial Intelligence*. San Francisco: Morgan Kaufmann Publishers Inc., pp. 1148–53.
- Vezzani, Federica, and Giorgio Maria Di Nunzio. 2019. (Not so) Elementary, my dear Watson! A different perspective on medical terminology. *Umanistica Digitale* 3: 59–75. [CrossRef]

- Vezzani, Federica, and Giorgio Maria Di Nunzio. 2020. Methodology for the Standardization of Terminological Resources. Design of TriMED Database to Support Multi-register Medical Communication. *Terminology. International Journal of Theoretical and Applied Issues in Specialized Communication* 26: 265–97. [\[CrossRef\]](#)
- Vezzani, Federica, Piccini Silvia, Costa Rute, and Giorgio Maria Di Nunzio. 2023. Le sens de la polysémie en terminologie: Propositions de représentation. In *En termes de polysémie. Sens et polysémie dans les domaines de spécialité*. Verona: Università di Verona, May 25–26.
- Vezzani, Federica. 2019. La technicité des termes: Le *v-tech* comme paramètre d'évaluation. In *Terminologie & Ontologie: Théories et Applications, Actes de la conférence TOTh 2019*. Chambéry: Presses Universitaires Savoie Mont Blanc, pp. 215–27.
- Vezzani, Federica. 2023. La connotation du vocabulaire somatique: Une étude de cas comparative bilingue en oncologie. *Meta: Journal des traducteurs/Meta: Translators' Journal* 68: 97–118. [\[CrossRef\]](#)
- Wierzbicka, Anna, and Вежбицкая Анна. 2021. 'Semantic Primitives', fifty years later. *Russian Journal of Linguistics* 25: 317–42. [\[CrossRef\]](#)
- Xie, Ruobing, Xingchi Yuan, Zhiyuan Liu, and Maosong Sun. 2017. Lexical Sememe Prediction via Word Embeddings and Matrix Factorization. In *Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence, Melbourne, Australia, 19–25 August*. Edited by Carles Sierra. Melbourne: International Joint Conferences on Artificial Intelligence, pp. 4200–6. [\[CrossRef\]](#)
- Ye, Yining, Fanchao Qi, Zhiyuan Liu, and Sun Maosong. 2022. Going "Deeper": Structured Sememe Prediction via Transformer with Tree Attention. In *Findings of the Association for Computational Linguistics: ACL 2022*. Edited by Smaranda Muresan, Preslav Nakov and Aline Villavicencio. Cedarville: Association for Computational Linguistics, pp. 128–38. [\[CrossRef\]](#)
- Zhang, Xiang, Ning Shi, Bradley Hauer, and Grzegorz Kondrak. 2023. Bridging the Gap Between BabelNet and HowNet: Unsupervised Sense Alignment and Sememe Prediction. In *Proceedings of the 17th Conference of the European Chapter of the Association for Computational Linguistics*. Edited by Andreas Vlachos and Isabelle Augenstein. Cedarville: Association for Computational Linguistics, pp. 2789–98. [\[CrossRef\]](#)

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.