

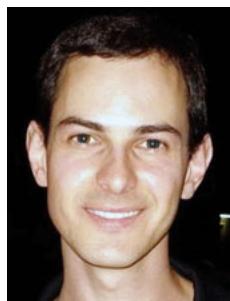
Knowledge Representation and Orthophemism: A Reflection Aiming to a Concept[†]

Suellen Oliveira Milani* and Fabio Assis Pinho**

* Doctoral Student, Graduate School of Information Science,
Sao Paulo State University–UNESP, Av. Hygino Muzzi Filho, 735 – Marília (17525-900),
São Paulo, Brazil, <suellen@marilia.unesp.br>
** Associate Professor, Department of Information Science, Federal University
of Pernambuco–UFPE, Av. da Arquitetura s/n – Recife (50740-550),
Pernambuco, Brazil, <fabiopinho@ufpe.br>



Sullen Oliveira Milani is a doctoral student at the Graduate School of Information Science, State University of São Paulo, Brazil, with collaborative calendar year at the School of Information Studies, University of Wisconsin-Milwaukee, with a grant by CAPES Foundation, Ministry of Education of Brazil. She has a master's degree in Information Science (2010) and bachelor's degree in Librarianship (2007), State University of São Paulo, Brazil. She researches the area of knowledge organization, more specifically under the theme of ethics in knowledge representation. Complete résumé available at <http://lattes.cnpq.br/2821128318835622>.



Fabio Assis Pinho is an Associate Professor in the Department of Information Science at the Federal University of Pernambuco, Recife-PE, Brazil. He received his master's in Information Science in 2006, and his PhD in Information Science in 2010, both from the São Paulo State University, Marília-SP, Brazil. His research interests lie in the broad area of knowledge organization and representation, especially the ethical issues in knowledge organization. His bibliographical production is registered at <http://lattes.cnpq.br/4220711855480007>.

Milani, Suellen Oliveira and Pinho, Fabio Assis. **Knowledge Representation and Orthophemism: A Reflection Aiming to a Concept.** *Knowledge Organization.* 39(5), 377-393. 29 references.

ABSTRACT: Characteristics of speech, especially figures of speech, are used by specific communities or domains, and, in this way, reflect their identities through their choice of vocabulary. This topic should be an object of study in the context of knowledge representation once it deals with different contexts of production of documents. This study aims to explore the dimensions of the concepts of euphemism, dysphemism, and orthophemism, focusing on the latter with the goal of extracting a concept which can be included in discussions about subject analysis and indexing. Euphemism is used as an alternative to a non-preferred expression or as an alternative to an offensive attribution—to avoid potential offense taken by the listener or by other persons, for instance, *pass away*. Dysphemism, on the other hand, is used by speakers to talk about people and things that frustrate and annoy them—their choice of language indicates disapproval and the topic is therefore denigrated, humiliated, or degraded, for instance, *kick the bucket*. While euphemism tries to make something sound better, dysphemism tries to make something sound worse. Orthophemism (Allan and Burridge 2006) is also used as an alternative to expressions, but it is a preferred, formal, and direct language of expression when representing an object or a situation, for instance, *die*. This paper suggests that the comprehension and use of such concepts could support the following issues: possible contributions from linguistics and terminology to subject analysis as demonstrated by Tálamo et al. (1992); decrease of polysemy and ambiguity of terms used to represent certain topics of documents; and construction and evaluation of indexing languages. The concept of orthophemism can also serve to support associative relationships in the context of subject analysis, indexing, and even information retrieval related to more specific requests.

Received 4 July 2012; Accepted 4 July 2012

[†] Coordination for the Improvement of Higher Level - or Education - Personnel - CAPES.

1.0 Introduction

The necessity for dialogue between knowledge representation, which often contains biases, and the inherent commitment of information systems to meet the demands of multiple user communities raises potential dilemmas to the indexer. In this sense, the main discussion occurs in a scenario of “well-doing” (Guimarães 2000) in an individual way (professional with her/himself) or in a social way (the professional in his/her relation to the users). In this context, this paper highlights the importance of mediating activities as the construction of thematic bridges within the context of production of documents and the context of acquisition and use of knowledge. These mediation activities can be understood as subject analysis (and more specifically classification and indexing), wherein the purpose is to isolate and label evidence from a recorded and socialized knowledge. For that, subject analysis relies on natural and artificial languages. So the understanding of how documents should be represented for effective retrieval and usage “is primarily a problem of language and meaning” (Mai 2001, 592). Tálamo et al. (1992) explains that “This fact imposes on the subject analysis procedures which are simultaneously flexible and rigorous.”

Subject analysis can be seen through two aspects: “theoretical, related to the determination of its own object of study, its functions and its methods; and pragmatic, which deals with rules in regards to the operation of documental systems. In this way, subject analysis must criticize its assumptions, procedures, and tools in order to advance both theoretically and practically” (Tálamo et al. 1992). While analyzing documents, indexers deal with some steps (neither in a narrower way nor as a rule) in order to reach the goal of “naming information” (Olson 2007). So, “[i]n this process, terms or notations assigned to reflect the document's subject are organized into a database such as a library catalog” (Olson 2001, 639). Mai (2001, 593) explains that “[i]t is useful, however, to operate with the idea of steps when analyzing the process [subject analysis], because breaking down the process into its individual parts will allow one to examine it in greater detail.”

In this sense, Tennis (2010) added two steps to the analysis of documents proposed by Mai (2001), one before and one after Mai's structure. So, subject analysis (which in the context of Tennis' study is re-

ferred to as micro-time stages) is compounded by: (1) Pre-analysis stage, (2) Document analysis processing, (3) Subject description, (4) Subject analysis, and (5) Indexing evaluation. The second step, document analysis processing, involves the comprehension of the textual structure of the document, and “has the objective to identify and select the concepts which represent the essence of a document” (Fujita 2003, 67). As Fujita (2003) highlights, a scientific document's textual super-structure contains: title (in its own language and in English), authorship, abstract (in its own language and in English), keywords (in its own language and in English), introduction (theme, goals, and theoretical approach), methodology (materials and methods), results, illustrations, discussion of the results, conclusions, and references. The third step, subject description, aims to create a cohesive formulation of the subject of the document in the language used by the author (Mai 2001), in other words, the formulation of an “indexing sentence,” which remains in natural language. Mai (2001) explains that what this paper calls an “indexing sentence” can be an unordered mass of mental impressions, phrases, terms, etc., but, before the next step, the indexer must select the concepts which could be useful to information retrieval and usage. The fourth step, subject analysis is the translation from an “indexing sentence” in natural language to a controlled vocabulary, or as Mai (2001, 594) says: “translating the product ... into a formal statement of the same thing, only this time in terms of the language of the appropriate subject access system.”

In the context of the first and last steps, the pre-analysis stage and the indexing evaluation, Tennis (2010, 226) explains that “the indexer must have some evidence and or intuition about how to settle on an interpretation, and this is in turn influenced by intention. These stages are also influenced by the materials present during indexing like the classification scheme, extant collection, and the like.” As results of these steps, the indexers create products such as: descriptors, subject headings, and indexes, abstracts, and notation of classification, with the aim to provide users' information retrieval, access, and usage. These products consist of “the completed subject entry from a given system that the indexer has finally chosen to represent the subject of the document” (Mai 2001, 594).

This paper will focus on the second step, document analysis processing, more specifically in the activity called “documental reading.” According to Cintra (1989), during word-to-word reading the indexer uses two reading movements:

- Bottom-up: from the part to the whole, ascending (analysis of the visual input), or deductive. The goal is to extract the meaning.
- Top-down: from the whole to the part, descending (previous knowledge and inference), or inductive. The goal is to assign meaning.

During the Bottom-up movement, the indexer may encounter some linguistic challenges, including figures of speech, and the way in which she/he deals with these terms¹ will influence assigning meaning, and can, in turn, exclude specific user communities.

It is known that terms themselves are complex. A good example was extracted from Tennis’ (2007, 91) analysis of the theme of eugenics. According to him, this concept “has affected a number of areas of science, social science, and philosophy. How words are used to present this concept affects the way it will be used by the indexer.” This paper will instead focus on the issues related to figures of speech and how they are presented within the literature used by specific user communities.

A figure of speech is “Any of the various ‘forms’ of expression deviating from the normal arrangement or use of words, which are adopted in order to give beauty, variety, or force to a composition” (Oxford English Dictionary 2012). Cintra et al. (2002) gives some examples:

- Ambiguity: a specific phenomenon of semantics where a word can have multiple meanings. One must recognize the concept of polysemy, which is a vocabulary phenomenon where a word can have plural meanings, and homonymy, where the same signifier refers to two different realities on separate occasions. Homonymy can refer to homophony—how words have identical sounds with different meanings—or homography—how words can have identical spellings, but different meanings;
- Synonymy: equivalence relation between at least two words;
- Hyponymy: relations between part/whole, genre/species; hierarchies;

- Euphemism: alleviates negative expressions through conceptual replacement.

For instance, when certain words and expressions are viewed as taboo, some changes in the language are motivated by creating new expressions (sometimes with a playful approach) and which can result in new meanings for old expressions. This prohibition of behavior applied to a particular group of people in a specific context can influence the linguistic behavior, resulting in the creation of euphemisms or dysphemisms, which, in turn, could hide prejudices.

In the Brazilian linguistic context, there are some examples: *chicken* to name a gay man, *monkey* to name a black person, both certainly used in a pejorative manner; moreover, linguistic misbehaviors are legally and socially punishable. In some way, there is an interaction among taboos, prejudices and the creation of euphemisms and dysphemisms (Radcliffe-Brown 1939). In this sense, Pinho (2010) and Milani and Guimarães (2011) demonstrate that there are some inadequacies related to some specific groups, such as gay men and women in Brazilian indexing languages including metaphors. In the context of taboos, the relationship between censorship and courtesy/discourtesy is categorized into groups of words and/or expressions resulting from this relationship in terms of a collective called “*n*-phemisms,” which is the union of the euphemism, dysphemism, and orthophemism (Allan and Burridge 2006).

This study aims to explore the literature of linguistic phenomena: euphemism, dysphemism, and orthophemism, focusing on the last with the goal to extract a concept which could offer assistance to the process of knowledge representation. This is an exploratory research whose starting point was the studies that have been coordinated by Professor José Augusto Chaves Guimarães at São Paulo State University (Marília, São Paulo, Brazil), and more specifically from the research developed by Pinho (2010) and Milani and Guimarães (2011). In this way, a bibliographic review is presented using the “close reading” technique to gain familiarity with the association between figures of speech and knowledge representation.

2.0 *N*-phemism: euphemism, dysphemism and orthophemism

Metaphorical words perform euphemistic and dysphemistic processes since their characteristic is, according to Guiraud (1975, 61), “that which designates unfamiliar beings or objects, and associates them to

others more well known.” Orthphemism, in turn, is the process by which metaphorical words perform in the strictest sense that which is linguistically ordinary in society’s many levels.

A euphemism, from the Greek *euphemismos* meaning “well said,” is used as an alternative to an unfavorable expression in order to avoid some characteristic loss, or as an alternative to an offensive attribution by the listener or another person; in other words, a euphemism is a figure of speech with the aim to replace an unpleasant word with another word to soften it without modifying the meaning. Euphemism is used to disguise unpleasant meanings, ruthless thoughts, or taboo words². In a certain way, euphemism portrays an extremely polite, even to the point of being forced, form of expression, which seeks to avoid a representation of reality that might be ruthless or unpleasant.

According to Casares (1959, 372), euphemism is the “way to say or suggest in disguised form, or with decorum, an idea whose straight and honest expression would be harsh, or would cause discomfort.” Because of that, Benveniste (1988, 340) attests that the term euphemism has a peculiar and paradoxical explanation at the same time, in that it threatens opposite senses, for instance, “say good augury words” and “avoid bad augury words.” On the other hand, a dysphemism is an expression with offensive connotations about the *denotatum*, that which refers to linguistic expression, or about the public, or both, and is therefore replaced by a euphemistic expression.

Chamizo Domínguez (2004, 45) explains that a specific word can be understood through contextual means, and, in this way, it is independent of the word itself—the speaker’s intentions are revealed through the context and usage of their words. This explanation is understood in the context of dysphemisms because the boundary between these concepts is restrictive. Considering that ambiguity is central to euphemism, candidate words must contain double meanings. In this way, a euphemism is unable to be replaced by another word and maintain the same cognitive, stylistic, and social effects due to the lack of synonyms in their natural language.

Euphemisms and dysphemisms are studied in the same way as metaphors. Because of this, they belong to conceptual networks, in other words, they organize and compose our thoughts and our actions. The primary goal of euphemisms is to label or to name an unpleasant object or its unpleasant effects. However, they also have secondary functions, such as the use of courtesy:

- to name someone, e.g., madam instead of miss;
- to dignify an occupation or labor, e.g., sanitary engineers instead of garbage collectors;
- to dignify a patient or a painful situation, e.g., seniors instead of old people;
- using political correctness, e.g., developing countries instead of Third World countries;
- to avoid ethnic or sexual harm, e.g., African-American instead of Negroes or Colored people;
- to name a taboo object or action, e.g., glistening instead of sweaty.

However, these examples demonstrate that a word used as a euphemism can work in a specific linguistic context, but can have a completely different meaning in another language. Although euphemism and dysphemism share several features with metaphors, they have different cognitive and social functions.

According to Allan and Burridge (2006, 31), dysphemism is the opposite of euphemism, and, in general, it is a taboo. Dysphemism is often motivated by fear and displeasure, as well as by hatred and contempt; in this way, speakers come up with dysphemism to talk about “people and things that frustrate and annoy them, that they disapprove of and wish to disparage, humiliate and degrade.” Dysphemisms are peculiar to groups addressing an opponent (or enemy). Dysphemistic expressions are profanity, name calling, and any kind of derogatory comment directed towards others with the aim to insult or hurt. This paper considers that dysphemism is a way to discharge emotion, for instance, when negative language is used in order to alleviate frustration or anger.

According to Allan and Burridge (2006, 2),

Orthphemism (Greek *ortho* - ‘proper, straight, normal’, cf. *orthodox*) is a term we have coined in order to account for direct or neutral expressions that are not sweet-sounding, evasive or overly polite (euphemistic), nor harsh, blunt or offensive (dysphemistic). For convenience, we have also created the collective term X-phemism to refer to the union set of euphemisms, orthphemisms and dysphemisms.

Orthphemism can be understood simply as direct expressions. Orthphemism is compound by: *ortho*-: “Forming words, chiefly scientific or technical, with the sense ‘straight, rectangular, upright, perpendicular’, or sometimes ‘right, correct, proper’” (Oxford English Dictionary 2012), and *-phem*: “a sign which functions as, or is equivalent to, a grammatical sen-

tence. ... an inflected or compound word, or a sequence of words, regarded as a unit having grammatical meaning rather than as a means of conveying sense and reference" (Oxford English Dictionary 2012). In summation, orthophemism introduces adjectives as proper and normal to the speaking act.

Because of that, orthophemism and euphemism are words or sentences used as an alternative to unfavorable expressions. These expressions are considered most suitable, in contrast to dysphemism, which can be considered as taboo. So, people who have enough linguistics comprehension to formulate euphemism and orthophemism have a social *savoir faire*, because "Just as we look after our own face (self-respect), we are expected to be considerate of, and look after, the face-wants of others" (Allan and Burridge 2006, 33).

In this context, orthophemism is more direct, formal, and literal than its euphemistic correspondent, which, in turn, is more colloquial, figurative, and indirect. A simple example of *n*-phemisms is related to the occasion of death, for instance: a euphemistic expression would be *pass away*, a dysphemic expression would be *kick the bucket*, and an orthophemic expression would be *die*.

Thus, *n*-phemisms are crossed synonymous variations; however, they are used in specific and alternative contexts because they share the same denotation while differing in connotation, which, in turn, arises from encyclopedic knowledge, beliefs, and prejudices about the contexts in which the word is commonly used. So if the connotative sense of *n*-phemisms arises from knowledge about certain prejudices, which operate in a specific context, the *n*-phemisms will inevitably contain negative characteristics. In some way, the latter statement suggests that there is an individual malleability operating under different circumstances, and settling her/his decision about the offensiveness or politeness included in her/his expression. Therefore, a euphemism can sound like a dysphemism to a specific person or vice versa.

However, there is a great possibility of euphemisms conveying or hiding some kind of prejudice caused by the offensive features assigned by speakers. If dysphemism on its own already conveys offensiveness and therefore negative connotations, and orthophemism covers direct and formal speech, so too can euphemism, with its soft-spoken approach, disseminate through this bias a hidden prejudice to specific individuals.

This explanation is supported by sociolinguistics, which reflects the many variations of language, and, because of this, language is not a static or rigid thing.

These variations occur in the language's structural or internal context as well as in the social or external context, and, in this way, this paper considers language to be a system that changes according to the changes in the social structure, whose recontextualization will occur in its internally. Sociolinguistic studies explore the relations between language and the speaker's social condition taking sociocultural factors into consideration (Prete 2000). Sociolinguistics considers the social importance of language and studies its variations. According to Mollica (2004, 9), "Sociolinguistics is one of the subcategories of linguistics studying language in usage within speaker communities, motivating an investigation which correlates linguistic and social aspects ... focusing on concrete linguistic usages, especially on those with heterogeneous features."

Hence, the indexer in the moment of professional performance, in other words during indexing, tries to organize information, as well as to label it, most often permeated by certain subjectivity. It is true that the indexer performing this task bases his or her decisions on a specific cultural and ideological context, and, in this way, is involved with linguistic issues that, in turn, are involved with *n*-phemisms. The indexer must identify in the context of recorded and socialized knowledge, to what extent the mentioned figures of speech are used before the moment that they are fed into a database, catalog, or even an indexing language.

Semantic changes aiming to fit a meaning are present in some examples, such as *queer*³, which can be understood defined as "a. Strange, odd, peculiar, eccentric. Also: of questionable character; suspicious, dubious" (Oxford English Dictionary 2012). However, "*queer*" is synonymous with "*gay*," which can designate a bias. Meanings related to *queer* can be bidirectional; on one hand, this word refers to a set of expressions reinforced into the homophobic vernacular, in this way pejorative, and, on the other hand, *queer* refers to an Anglo-Saxon definition, whose meaning is related to something strange, unusual, or peculiar.

It is interesting to observe that when the indexer does use the word *queer* in a subject heading, she/he will do it according to literary warrant, but while recognizing that this word can have other meanings. The indexer will thus deal with Terminology, which is dedicated to observation of the usage of terms and their relation to scientific knowledge. In this way, Terminology implies theoretical elements and practical principles, which are able to guide searches, selec-

tion, and arrangement of terms belonging to specific knowledge domains.

Thus, this paper understands that in order to create "surrogates of knowledge" (Olson 2002), studies about figures of speech existing in a knowledge domain start with the identification of their linguistic system and their lexical universe, resulting in a terminological set which will be used in a specific domain; in other words, they are word-cultural sets. This notable identification characterizes euphemistic, dysphemistic, and orthophemic words.

3.0 Discussion

Milani and Guimarães (2011) presented an examination about how women are presented in the following Brazilian indexing languages: Subject Terminology of the National Library, University of São Paulo Subject Headings, Brazilian Senate Subject Headings, and Law Decimal Classification. The results showed signs of biases, which enabled the proposition of guidelines that may contribute to minimizing these biases and open the way for further discussions.

Considering that indexing languages are controlled vocabularies, or non-natural languages, or artificial languages constructed in a way to establish preferred descriptors and their relationships in order to complete the subject analysis, it is possible to say that the presence of figures of speech is not common in indexing languages.

Cintra et al. (2002, 35) explains that:

the system of relationships into Indexing Languages is not dynamic, as opposed to natural language, as well as their mechanisms of articulation, which are extremely precarious in view of those existing in language in general. On the contrary, elements of this specific language are selected from determined universes, and its system of relationships is constructed, necessitating the existence of explicit rules to use it. Because of that, the Indexing Languages are constructed languages.

Milani and Guimarães (2011) did not find examples of figures of speech within Brazilian indexing languages, but this situation would be very different if the authors had analyzed the literature related to women's issues. Considering that librarians perform subject analysis based on the context of the standard which they use, it is important to present some examples extracted from Olson (2002).

In *The Power to Name*, Olson (2002) analyzes the application of *Library of Congress Subject Headings* and *Dewey Decimal Classification* to: a) three books with: "specific and concrete topics (concerned with aspects of race, class and gender)," b) two books with "complex, but specific and theoretical topics," and c) two books and their treatment in terms of exhaustivity, in other words, one heading was originally assigned to the first book, and many headings were originally assigned to the last book.

With the goal to provide examples about figures of speech in the context of women literature, this paper analyzed the titles and Olson's conclusions about those books. The figures of speech assigned by the authors of this paper to each book are:

- *Chain Her by One Foot: the Subjugation of Native Women in Seventeenth-century New France* (Karen Anderson, 1991) - Synecdoche;
- *Secretaries Talk: Sexuality, Power, and Work* (Rosemary Pringle, 1989) - Synecdoche;
- *Maid in the U.S.A.* (Mary Romero, 1992) - Ambiguity, more specifically Homophony;
- *Talking Back: Thinking Feminist, Thinking Black* (bell hooks, 1989) - Metaphor;
- *When the Moon Waxes Red: Representation, Gender, and Cultural Politics* (Trinh T. Minh-ha, 1991) - Allusion;
- *Immortal, invisible: Lesbians and the Moving Image* (Tamsin Wilton, 1995) - Allusion; and,
- *The Myth of Aunt Jemima: Representations of Race and Region* (Diane Roberts, 1994) - Metonymy.

These examples work as illustrations to the concepts presented in this paper, but it is important to note in Olson's dissertation how cataloguers dealt with those complex themes and which headings and classification numbers (from *Library of Congress Subject Headings* and from *Dewey Decimal Classification*) were assigned to each book. It is important to highlight that sometimes the subtitles can clarify the scope of the book—other times they just present the context in which the book is presented, or they did not help in any way.

Besides these examples, it is important to present some examples extracted from the literature of male homosexuality, such as: Bareback, Bears, Coming out, Don't ask, don't tell, Chicken, Dude, Fag hag, Cruising, and Dogging. It was noted that these words are metaphorical and euphemistic, and they are used for individuals that belong to this domain. Orrico (2001) argues that the use of figures of speech can in-

crease the feedback accuracy for specific information requests, and, for that, she proposes a meta-filter, whose guiding axes are semantic because she understands that human beings represent themselves and their world using metaphors. The author therefore bases her ideas on two fundamental spheres: the discourse sphere and identity sphere, in other words, groups can construct their identities into discourse and through figures of speech, so metaphors can establish or maintain significant relations.

Orrico (2001, iv) clarifies that

assuming metaphor as a semantic core that is able to retrieve information implies admitting that a set of metaphorical representations as well as their linguistic manifestations compose the semantic-conceptual universe of a specific discursive community, keeping in mind that it is through discourses' interactive change that meanings are constructed.

In a homosexual context, discourses are permeated by figures of speech—euphemisms and dysphemisms—which are appropriated by individuals to both mischaracterize and recharacterize them and adopt a new meaning. To disregard this statement is to disregard the identity construction of this group—and therefore the aspects of these figures of speech must be considered by librarians engaging in subject analysis and indexing. Appropriating these terms means that, in the context of marginalized groups, marginalization overcomes through a new interpretation of the term.

The meta-filter proposed by Orrico (2001, 125) seeks to solve the problem involving the representation of metaphors by taking into account the ambiguity contained within language, users' necessities, and topics of documents. Such a proposal has the following fundamental conditions: Essence, Function, and Mode. Essence corresponds to the key idea as ontological metaphor, followed by a set of metaphors which represent the Function, and receive direct influence from Essence, and, lastly, a function of realization that is the Mode.

Orrico (2001, 134) recognizes that

the use of metaphors would work in a meta-structural scheme, which would guide the organization of other schemes: the item layer - which establishes a direct relation between the term and documental inscription; the meta-concept layer - which brings together categories from a set of related elements; and the domain

layer - which bounds the field of significance where the metaphor is located.

In this sense, it is possible to say that figures of speech related to specific discursive communities or knowledge domains are important and can be included in knowledge representations performed in the context of such topics. Because of this, what this paper ultimately proposes is that representations arising from figures of speech regarding certain domains, as a way of human beings identifying the world, comply with some rules attested by human beings belonging to this area. These representations 'comply' with conformities that can be used to organize the concepts of this domain (Orrico 2001).

Thus, this paper reinforces that orthophemism establishes an important connection among word meanings beyond the use of figures of speech, since orthophemism establishes important associative relationships in the scope of knowledge representation. *The Guidelines for the Construction, Format, and Management of Monolingual Controlled Vocabularies* (2005, 51) clarifies that:

This relationship covers associations between terms that are neither equivalent nor hierarchical, yet the terms are semantically or conceptually associated to such an extent that the link between them *should* be made explicit in the controlled vocabulary, on the grounds that it *may* suggest additional terms for use in indexing or retrieval. The most common associative relationship used in thesauri is symmetrical and is generally indicated by the abbreviation RT (related term), but more semantically refined associations may also be developed to capture both symmetric and asymmetric associations.

The *Guidelines* highlight that the associative relationship can link terms under certain circumstances, primarily terms which belong to different hierarchies. In this way, terms belonging to the same hierarchy can also have relationships between: Overlapping Sibling Terms (e.g., ships and boats), or Derivational Relationships (e.g., mules: Broader Term equines, RT donkeys, RT horses). It is not necessary to link Mutually Exclusive Sibling Terms (e.g., roses and daffodils).

Terms belonging to different hierarchies can be linked according to the following associative relationships: a) Process/Agent, b) Process/Counteragent, c) Action/Property, d) Action/Product, e) Action/Target, f) Cause/Effect, g) Concept or Ob-

ject/Property, h) Concept or Object/Origins, i) Concept or Object/Units or Mechanisms of Measurement, j) Raw Material/Product, l) Discipline or Field/Object or Practitioner, m) Noun is not a true broader term (e.g., fishes and fossil fishes), or Antonyms.

Considering that the *Guidelines* (2005, 57) point out that "In certain controlled vocabularies, it *may* be considered desirable to refine Related Term references in order to make the nature of the relationships explicit. Codes for such relationship indicators and their reciprocals *may* be developed locally," figures of speech, or more specifically euphemisms, dysphemisms, or orthphemisms could occupy positions as *overlapping sibling terms*.

A specific collection or a specific library can also have its own code related to euphemisms, dysphemisms, and orthphemisms, whose relationship indicators could be established according to literary warrant (Barité Roqueta 2011) or following the principles of cultural warrant (Beghtol 2002). The measure of how fruitful and/or stable these associative relationships can be, will demand ethical decision-making (Beghtol 2005) because figures of speech are dynamic and thus their essential characteristics are related to constant change and replacement of concepts.

4.0 Conclusion

This paper presents the concepts of euphemism, dysphemism, and orthphemism as objects of study of subject analysis and indexing. Figures of speech are present in documental production and must be taken into account during subject analysis, primarily in relation to specific user communities or knowledge domains.

This first incursion into the literature demonstrates that the comprehension and use of such linguistic phenomena could support the following issues:

- Possible contributions from linguistics and terminology to subject analysis as demonstrated by Tálamo et al. (1992);
- Decrease of polysemy and ambiguity of terms used to represent certain topics of documents; and,
- Construction and evaluation of indexing languages.

This paper also suggests that the second step of subject analysis proposed by Mai (2001), document analysis processing, is a favorable moment to reflect on the presence of figures of speech and their possi-

ble impacts, as well as indexing evaluation (Tennis 2010), where one's indexing decisions can be re-evaluated. These statements must be understood as a starting point for further research, considering that the concepts of euphemism, dysphemism, and orthphemism, from their linguistic approach, specify the context of terms used for specific discursive communities or knowledge domains. This factor is relevant to the accuracy in the processes of knowledge representation.

Orthphemism is a word or expression used as an alternative to unfavorable expressions. In the scope of subject analysis and indexing, the use of euphemisms and orthphemisms can support information retrieval related to more specific requests, providing signs of cultural warrant to the information system. The analysis of expressions of this nature should be done by the indexer during subject analysis, before the insertion of certain candidate words or expressions to indexing documents. The insertion of euphemisms and orthphemisms in indexing practice and even into indexing languages to establish semantic relations among terms, can clarify to users how a certain domain is organized, and can even encourage the preservation of that domain.

If euphemisms and dysphemisms can be loaded with prejudices, orthphemism can in turn be related to extremes of political correctness. Considering that language is dynamic, this paper emphasizes the importance of ethical practices by the indexer when establishing labels and semantic relations among indexing terms. Once analyzed, orthphemism can be understood as an ally in the context of associative relationships and information retrieval. Orthphemism, which is formal and direct, is a favorable language of expression to represent an object or a human being's action. This is the clear and precise way to express a concept without the need to hide one's intentions or use false sincerity. Orthphemism also has cultural features, which can be linked to knowledge representation with a low possibility of resulting in harms or prejudices. Thus, it suggests that studies about orthphemism, as an axiologically neutral term, or a term which is strictly referential or descriptive, could yield fruitful contributions to the teaching and research of courses related to subject analysis and indexing.

Notes

1. Tálamo et al. (1992) brings forward the importance of terminology to subject analysis, and the

- association with terminological tools and indexing languages, which would improve the methodological accuracy of indexing and retrieval processes. The authors explain that "Unlike what occurs in indexing languages, the definition of terminological unities - called terms - occurs from the discourse, and not from the language. This means that the term has a specific and relational concept, whereas the unities of the indexing languages [or descriptors] are defined as free predicates."
2. Taboo "is the prohibition of certain actions or objects based on religious reasons, conventions, social attitudes, or other prejudices while the linguistic taboo is the word whose use should be avoided, generally for social, political, sexual, or religious reasons" (Moreno Fernández 1992, 201).
3. Gorman-Murray et al. (2010, 99) highlights that the concept *queer* is used to represent lesbian and gay lives as they exist outside of normative heterosexuality, and how communication works across and subvert subjective binaries in research relationships, such as insider/outsider. "The notion of queer asserts the multiplicity and fluidity of sexual subjectivities."
- ## References
- Allan, Keith, and Burridge, Kate. 2006. *Forbidden words: taboo and the censoring of language*. New York: Cambridge University Press.
- Barité Roqueta, Mario. 2011. "La garantía literaria como herramienta de revisión de sistemas de organización del conocimiento: modelo y aplicación." PhD diss., University of Granada.
- Beghtol, Clare. 2002. A proposed ethical warrant for global knowledge representation and organization systems. *Journal of Documentation* 58: 507-32.
- Beghtol, Clare. 2005. Ethical decision-making for knowledge representation and organization systems for global use. *Journal of the American society for information science and technology* 56: 903-12.
- Benveniste, Emile. 1988. *Problemas de lingüística geral I*. Campinas: Editora da Universidade Estadual de Campinas.
- Casares, Julio. 1959. *Diccionario ideológico de la lengua española: desde la idea a la palabra; desde la palabra a la idea*. Barcelona: Gustavo Gili.
- Chamizo Domínguez, Pedro J. 2004. La función social y cognitiva del eufemismo y del disfemismo. *Panace@ 5n15*: 45-51.
- Cintra, Anna Maria Marques. 1989. Estratégias de leitura em documentação. In Smit, Johanna W. et al. eds., *Análise documentária: a análise da síntese*. Brasília: IBICT, pp. 29-37.
- Cintra, Anna Maria Marques et al. 2002. *Para entender as linguagens documentárias*. São Paulo: Editora Polis.
- Fujita, Mariângela Spotti Lopes. 2003. A identificação de conceitos no processo de análise de assunto para indexação. *Revista digital de biblioteconomia e ciência da informação* 1n1: 60-90.
- Gorman-Murray, Andrew, Johnston, Lynda, and Waitt, Gordon. 2010. Queer(ing) communication in research relationships: a conversation about subjectivities, methodologies and ethics. In Browne, Kath, and Nash, Catherine J. eds., *Queer methods and methodologies: intersecting queer theory and social science research*. Farnham, England: Ashgate, pp. 97-112.
- National Information Standards Organization. 2005. Guidelines for the Construction, Format, and Management of Monolingual Controlled Vocabularies. Bethesda, MD: NISO Press.
- Guimarães, José Augusto Chaves. 2000. O profissional da informação sob o prisma de sua formação. In Valentim, Marta Lígia Pomim, ed. *Profissionais da informação: formação, perfil e atuação profissional*. São Paulo: Polis, pp. 53-70.
- Guiraud, Pierre. 1975. *A semântica*. São Paulo: Difel.
- Mai, Jens-Erik. 2001. Semiotics and indexing: an analysis of the subject indexing process. *Journal of documentation* 57: 591-622.
- Milani, Suellen Oliveira, and Guimarães, José Augusto Chaves. 2011. Biases in knowledge representation: an analysis of the feminine domain in Brazilian indexing languages. In Smiraglia, Richard P. ed. *Proceedings of the Third North American Symposium on Knowledge Organization, June 16-17, 2011, Toronto, Canada*. Available <http://www.ideals.illinois.edu/handle/2142/27734>.
- Mollica, Maria Cecília. 2004. Fundamentação teórica: conceituação e delimitação. In Mollica, Maria Cecília, and Braga, Maria Luíza eds., *Introdução à sociolinguística: o tratamento da variação*. São Paulo: Contexto, pp. 9-14.
- Moreno Fernández, Francisco. 1992. *Principios de sociolinguística y sociología del lenguaje*. Barcelona: Ariel.
- Olson, Hope A. 2001. The power to name: representation in library catalogs. *Signs: journal of women in culture and society* 26: 639-68.
- Olson, Hope A. 2002. *The power to name: locating the limits of subject representation in libraries*. Dordrecht: Kluwer Academic Publisher.

- Olson, Hope A. 2007. How we construct subjects: a feminist analysis. *Library trends* 56n2: 509-41.
- Orrico, Evelyn Goyannes Dill. 2001. "Binômio linguística-ciência da informação: abordagem teórica para elaboração de metafiltro de recuperação da informação." PhD diss., Instituto Brasileiro de Informação em Ciência e Tecnologia / Universidade Federal do Rio de Janeiro.
- Oxford English Dictionary. 2012. Oxford: Oxford University Press. Available <http://www.oed.com.ezproxy.lib.uwm.edu/>.
- Pinho, Fabio Assis. 2010. "Ethical issues in knowledge representation on themes in male homosexuality: an analysis of accuracy in Brazilian indexing languages." PhD diss., São Paulo State University.
- Preti, Dino. 2000. *Sociolinguística: os níveis da fala: um estudo sociolinguístico do diálogo na literatura brasileira*. São Paulo: EDUSP.
- Radcliffe-Brown, Alfred Reginald. 1939. *Taboo*. Cambridge, England: Cambridge University Press.
- Tálamo, Maria de Fátima G. M., Lara, Marilda Lopes Ginez de, Kobashi, Nair Yumiko, and Keiko O. F. Amaro, Regina. 1992. A interface análise documentária, linguística documental e terminologia. In *III Simpósio Iberoamericano de Terminología*, San Millán de la Cogolla, España, 1-5 diciembre, 1992. Available <http://www.riterm.net/actes/3simposio/talamo.htm>.
- Tennis, Joseph T. 2007. Scheme versioning in the semantic web. *Cataloging & classification quarterly* 43n3/4: 85-104.
- Tennis, Joseph T. 2010. Measured time imposing a temporal metric to classificatory structures. In Gnoli, Claudio, and Mazzocchi, Fulvio, eds. *Paradigms and conceptual systems in knowledge organization. Advances in knowledge organization* 12. Würzburg: Ergon-Verlag, pp. 223-28.