

A new methodology for conceptual metaphor detection and formulation in corpora: A case study on a mental health corpus

Marta Coll-Florit
Universitat Oberta de Catalunya

Salvador Climent
Universitat Oberta de Catalunya

Abstract

We describe a new methodology for conceptual metaphor detection and formulation in corpora, developed within the framework of the MOMENT project for analysing mental health metaphors. We critically review state-of-the-art methods for metaphor identification in texts, highlighting their main drawbacks for metaphor analysis in large corpora, mainly practical applicability and analytical subjectivity. Our method aims at mitigating existing drawbacks on the basis of applying the following principles: (i) working hypothesis formulation and verification at the metaphorical expression detection stage; (ii) partial use of standard methods for metaphorical focus identification; (iii) use of external expert knowledge in the form of more extensive use of dictionaries and the additional use of metaphor compendia; and (iv) the implementation of strategies for conceptual metaphor formulation, including domain formulation at two levels of generalization. Satisfactory reliability test results were obtained when we tested our method for inter-annotator agreement regarding metaphor detection and formulation using texts about mental disorders as a test corpus.

Keywords: conceptual metaphor, corpus annotation, metaphor identification methods, conceptual metaphor formulation, mental health

1 Introduction

This paper presents a new methodology for annotating conceptual metaphors (CMs) in corpora. The conceptual metaphor theory, developed by Lakoff & Johnson (1980), posits that metaphor is so pervasive in ordinary daily life that our conceptual system, in terms of which we both think and act, is fundamentally metaphorical in nature (Lakoff & Johnson 1980: 3). According

to the theory, CMs are defined as a cognitive process by which a set of mappings is established between source and target conceptual-semantic domains, where the source domain (SD) is more concrete and more directly related to experience, in contrast with the target domain (TD), which is more abstract and diffuse and lacks clear delineation (Kövecses 2002: 20). As one example, it is very common to refer to ideas in relation to food (1). Note that any given CM may be linguistically instantiated in multiple metaphorical expressions (MEs), i. e., MEs are specific linguistic cases of a CM.

- (1) CM: IDEAS ARE FOOD
 ME: a. That's food for thought.
 b. I just can't swallow that claim.
 c. His idea was half-baked.

Another key notion in conceptual metaphor theory is associated with hiding and highlighting. According to Lakoff & Johnson (1980: 10), “the very systematicity that allows us to comprehend one aspect of a concept in terms of another [...] will necessarily hide other aspects of the concept”. Consequently, the use of different CMs to refer to a single domain may reveal different conceptualizations or ways of understanding the domain. This is why studying CMs is useful to detect implicit ideas and assumptions in discourse.

Discourse analysis based on conceptual metaphor theory has been implemented in numerous fields, including advertising, medicine, architecture, economics and religion (Soriano 2012: 117). Our research is conducted in the framework of a project titled *MOMENT: Metaphors of Severe Mental Disorders* (Coll-Florit et al. 2018), whose main focus is the application of conceptual metaphor theory to the mental health field. The primary goal of the MOMENT project is to contribute to a better understanding of severe mental disorders by analysing the discourse of people with those disorders and the discourse of mental health professionals. More specifically, first-person accounts by both groups, produced in Spanish and published on the Internet (in blogs, forums, etc), are analysed with the aim of detecting (1) the kind of metaphors used by these two groups when referring to mental disorders and related experiences; and (2) the kind of frames or interpretative discourses built on the use of those metaphors.

The MOMENT project involves the annotation of a large and heterogeneous corpus of mental health discourses, where the formulation of CMs – i. e., the identification and formulation of SDs and TDs for CMs

from MEs in texts – is a key component of the analysis. However, metaphor annotation in corpora is a challenging task. For example, the volume of a corpus imposes certain restrictions on manual annotation, the analysts may not always be first level experts, there is likely to be time constraints for the analysis of MEs and the corpus may be very heterogeneous, with a variety of subjects, sources and textual genres. An additional complication in our case is the adoption of a loose approach in the detection of TDs, which only need to exhibit a relationship to some aspect of mental disorders (the life of the affected people, the symptoms, the related emotions, the medication and professional intervention, etc).

To reduce the problems of time and data diversity, in a second stage of the project semiautomatic strategies of analysis will be adopted, in line with those proposed by Stefanowitsch (2006), Ogarkova & Soriano (2014) and Semino et al. (2018). However, manual annotation of a well-balanced subcorpus of equivalent texts for each group under analysis is first necessary in order to identify the prevailing set of CMs. The methodology proposed here describes this manual annotation procedure.

In this paper we focus on CMs. Nevertheless, it is important to note that two figures directly related to CMs are also annotated in our project: (1) conceptual metonymies, which represent an entity or concept in terms of another (like metaphors), but in this case relating two entities in the same conceptual domain, and (2) metaphorical similes, which are overt figurative comparisons signalled by a comparative marker.¹ This paper mainly discusses CMs, but the section describing the annotation method also briefly presents the analytical steps concerning conceptual metonymies and metaphorical similes.

Below we describe state-of-the-art methods for metaphor detection and formulation, their main drawbacks and the approach adopted in the MOMENT project to overcome those drawbacks (section 2). Next we describe the implementation of our manual annotation method, including details about the annotation steps and the documents that accompany the annotation guidelines (section 3). Finally, we provide the results of our reliability tests (section 4), briefly discuss some variables to take into account regarding the reproducibility of our method (section 5) and present our overall conclusions (section 6).

¹ Collectively we will refer to all three figures as “figurative language” and “figurative conceptualization”, while drawing the distinction between these concepts when necessary.

2 Existing methods for identifying metaphorical expressions and conceptual metaphors

The state-of-the-art methodology for ME detection in texts is the metaphor identification procedure (MIP, Pragglejaz Group 2007), subsequently refined and extended by Steen et al. (2010b) as MIPVU (where VU stands for Vrije University in Amsterdam). Steen's five-step procedure (1999; 2007) is the reference method for CMs accounting. Although it is not explicitly stated by their authors, these methods best apply to metaphorical uses of lexical words; thus they are seldom applied to detect metaphorical meanings expressed by grammatical elements.

The Pragglejaz Group (2007), in fact, focuses on identifying metaphorically used words (MUWs). The procedure is based on the distinction for all the words in the text between their meaning in that context and their so-called basic meaning. If the contextual meaning can be understood by comparison to the basic meaning, then a metaphoric use of the word in the text is determined. A reference dictionary is consulted to minimize errors and inconsistencies between analysts. Which of the dictionary meanings corresponds to the basic meaning is established by the analyst according to the following brief instruction regarding basic meanings, which "tend to be more concrete; what they evoke is easier to imagine, see, hear, feel, smell, and taste; related to bodily action; more precise (as opposed to vague); and historically older" (Pragglejaz Group 2007: 3). Note that the last condition has been ruled out by Steen et al. (2010a: 183) on the basis that older meanings are not necessarily more concrete.

By way of example, *riding* in (2)² corresponds to word sense 4 in the online version of the Macmillan Dictionary:³ "to float, or to appear to float, on water or in the air", while its basic meaning corresponds to word sense 1: "to sit on an animal, especially a horse, and control its movements as it moves along". The metaphor is meaningful in that, just as a rider controls the movements of a horse in movement, so too do the mermaids control the movements of the waves.

² The original source of this example is Steen (1999). In this and subsequent examples, the alleged MUW or metaphorical focus is boldfaced.

³ The Pragglejaz Group (2007) used the *Macmillan English Dictionary for Advanced Learners* (Rundell & Fox 2002) for their work. The online version used by us can be found at <https://www.macmillandictionary.com/> (accessed 2019-03-15).

- (2) I have seen the mermaids **riding** seawards on the waves.

MIP and MIPVU are not intended to identify the CMs underlying MUWs, although the Pragglez Group (2007: 34) does point out that the part of the procedure associated with defining basic meaning may be profitably used “to identify the source and target domains underlying metaphorical words in context”. In contrast, Steen (1999; 2007), after a first step devoted to establishing the so-called “focus” of a ME (a concept equivalent to the Pragglez Group’s MUW), does identify CMs, described in terms of propositional logic (see (7) in § 2.2.2).

In spite of their importance and widespread recognition, the Steen (1999; 2007) and Pragglez Group (2007) methodologies have several drawbacks, primarily practical applicability and subjectivity in the analyses. These will be discussed in more detail in the following sections.

2.1 The problem of time

One obvious problem with MIP and MIPVU is their application to the large-scale manual analysis of texts. The literature does not report information on the average time spent per analyst and the volume of text analysed, but it seems clear that manual analysis is not feasible for a corpus of tens of thousands of words, for which every word has to be looked up in a dictionary and then analysed for contextual and basic meanings. This problem has been tackled by some authors working with large corpora by narrowing down the range of domains to be analysed. Thus, Ogarkova & Soriano (2014) pre-selected domain-representative keywords to extract concordances using an extension of Stefanowitsch’s (2006) methodology, whereas Semino et al. (2018) combined manual with automatic semantic text labelling to filter out concordances and then selected those tagged with the most promising categories.

2.2 The problem of subjectivity

Several indeterminacy problems in metaphor analysis have been detected and analysed by Heywood et al. (2002), Semino et al. (2004), Valenzuela & Soriano (2005) and Geeraerts (2010), among others. Broadly speaking, two main difficulties are encountered: determining whether an expression is literal or metaphorical and, if metaphorical, formulating the underlying CM in terms of the appropriate SD and TD.

2.2.1 Identifying metaphorical expressions

Geeraerts (2010: 250) points out that the MIP fails to avoid the traditional difficulty of identifying semantic phenomena, namely, that “the possibility of understanding one reading in comparison with another depends on the interpreter’s ability to see the analogy – a highly subjective skill”.

Semino et al. (2004: 1277–1280) also point out that metaphoricity may be a matter of degree, as the boundary between the literal and the metaphorical is often fuzzy, as illustrated by the way the spread of cancer is talked about in terms of movement: “Although the development of cancer frequently involves the literal movement of cancerous cells inside the body (3a), the use of lexis to do with motion in our corpus often appears to be metaphorical (3b, 3c), since ‘coming back’ and ‘travelled’ do not apply literally to tumour or cancer” – we understand that because such verbs require animate agents. In any case, establishing their meaning in the text (possibly by comparison to an allegedly different literal meaning) is potentially subjective.

- (3)
- a. the way it gets there is through the blood stream.
 - b. chemotherapy can reduce the chances of things **coming back**.
 - c. it hasn’t **travelled** any more.

It is apparent that decisions concerning the literal meaning of a word can determine the metaphoricity of the phrase or sentence where that word is used. Regarding the word *freedom* in (4), for instance, Heywood et al. (2002: 46–47) suggest that its use “could be seen as metaphorical if one decided that the concept it evokes relates directly to domains such as slavery and imprisonment”, whereas “a more general interpretation of the concept of freedom as relating to the ability to choose” would point to a literal meaning.

- (4) they had so kindly offered **freedom**.

Note that, before the advent of the MIP, the focus of the debate regarding CM identification was discriminating between the figurative and literal meanings of words and expressions. However, the fact that the MIP draws a distinction between figurative and basic, rather than literal, meanings greatly simplifies the problem, as the MIP offers instructions for establishing basic meaning, whereas no instructions are available for establishing literal meanings. But, on the other hand, the MIP favours figurative readings since, while the distinction between the figurative and the literal meaning of a word is not always clear, the distinction between a figurative and a basic meaning is more straightforward.

In (4), for instance, the authors eventually decided that *freedom*, in the text, was not metaphorical since it was related to “the ability to choose what one wants to do without constraints from others”, which they considered to be the literal meaning of the word. However, an analysis in MIP terms would lean the analyst towards a metaphorical reading, since a so-called basic meaning can be found in dictionaries, e. g. “a situation where you are able to go where you want because you are not in prison” (Rundell & Fox 2002). As a consequence, following the MIP, the use of freedom in (4) would be considered metaphorical, as it contrasts with such a basic meaning.

As Geeraerts (2010: 207ff.) points out, an important source of indeterminacy is the polysemous structure of lexical units, and, at the very least, we need to consider the possible existence of dead metaphors, i. e., expressions that may be metaphorical from a diachronic point of view but have lost their metaphorical motivation for the average contemporary user. An example is (5a), which may not need to be accounted for by means of the metaphor A MOUNTAIN IS A PERSON. Otherwise, “there would just be an extension of the semasiological structure of foot, whereas the meaning of mountain could be left for what it is” (Geeraerts 2010: 208). Similar polysemy-related challenges for analysts arise when meanings originating in metaphorical extension are more frequent in contemporary language than those which could be reasonably considered literal. Semino et al. (2004: 1284–1285) report the case of *erupt* (5b) and *eruption*. Entries with those words in corpora of British English primarily reflect activities, emotions and different kinds of entities, and only reflect volcanic activity in 38% of the entries.

- (5) a. the **foot** of the mountain.
 b. something is gonna suddenly **erupt** and it’s all going to be all over.

Some cases of indeterminacy are seemingly unsolvable even considering the semasiological structure of lexical units. Heywood et al. (2002: 46) exemplify the case with (6), which seems to work both literally and metaphorically. While the expression in context can be analysed literally as being to do with the current physical location of *they*, on the other hand, “the possibility of a mapping from the domain of location to the domain of human activities in general means that the question could be to do with what has become of them”. In such cases, the decision on metaphoricity will rely on a very fine-grained interpretation of the co-text.

(6) **Where** had all they **gone**?

While aware of the complexity of these and similar problems, Ogarkova & Soriano (2014) and Semino et al. (2018) make flexible use of the MIP/MIPVU for ME identification in their analyses of large corpora, i. e., they draw on the contrast between contextual and basic meanings.

2.2.2 Formulating conceptual metaphors

The second main problem concerning subjectivity in metaphor analysis is domain labelling. Ding (2011: 72) notes that, while the MIP is a well-established procedure for identifying MUWs, it does not address CM formulation, and, furthermore, that the main problem with Steen's (1999; 2007) method is that it does not offer instructions to help in determining CM domains and correspondences.

As mentioned previously, Steen (1999; 2007) describes linguistic metaphors in terms of propositional logic, which associates concepts expressed in a text with underlying concepts. Non-expressed concepts are initially variables and then are instantiated. Steen (1999: 67) makes the metaphor underlying (7a) explicit by first formulating (7b) which, in turn, is transformed into (7c). The interesting aspect is that the instantiation of F , y and y' , i. e., identification of the concepts metaphorically underlying those present in the text, relies on the analyst's intuition or linguistic and psychological knowledge – in other words, the interpretation is subjective. The only instruction is that “prototypical or default knowledge about the source domain” (Steen 1999: 71) has to be activated.

- (7) a. I have seen the mermaids **riding** seawards on the waves.
 b. $F(\text{MERMAIDS}, \text{WAVES}) = \text{RIDE-ON}(y, y')$
 c. $\text{FLOAT}(\text{MERMAIDS}, \text{WAVES}) = \text{RIDE-ON}(\text{JOCKEY}, \text{HORSE})$

Steen's (1999) method, furthermore, does not allow clear labelling of the possible CM underlying (7a) as, by splitting the metaphoricity into three correspondences, TO FLOAT IS TO RIDE-ON, A MERMAID IS A JOCKEY and A WAVE IS A HORSE, it offers three prospects for concept labelling. Choosing labels for concepts or domains from among several possibilities is one of the most common problems in CM annotation, i. e., for a given scenario, the analyst might choose either a formulation based on the event or a formulation based on one of its arguments (Semino et al. 2004: 1276, 1281).

However, it is not only the internal structure of a certain frame, such as *to ride on horses* in (7), that can yield alternative domains for the metaphor. Domains may originate in entirely different frames, as Steen (1999: 71) and Geeraerts (2010: 207ff.) have pointed out. The latter author highlights *WAR* and *GAME* as alternative domains in cases like (8a), as *win* is commonly associated in the literature with the formulation *ARGUMENT IS WAR*. But the expression could be also perfectly consistent with *ARGUMENT IS GAME-PLAYING*, as in (8b). As Semino et al. (2004: 1284ff) have pointed out, domain selection can be biased by knowledge of extant conventional formulations.

- (8) a. I **won** the argument.
 b. lay one's cards on the table.

Ogarkova & Soriano (2014) introduced an interesting innovation by formulating CM in terms of two levels of generalization. In this way, they properly account for the theoretical distinction between generic-level and specific-level metaphors (Lakoff & Turner 1989: 80–81). In this approach, there is an inherited hierarchical structure among metaphors so that underspecified generic CMs pass on their structures to specific-level CMs. For example, *LIFE IS A JOURNEY* and *LOVE IS A JOURNEY* are specific-level metaphors of the generic *LONG-TERM PURPOSEFUL ACTIVITIES ARE JOURNEYS*. Likewise, Semino et al. (2004: 1291), while not unfolding domain formulation in levels, opt for mappings that most closely correspond to linguistic expressions.

2.3 Summary

To sum up, we list the following main problems with standard methods for identifying MEs and CMs in analyses of large corpora, of direct relevance, moreover, to our *MOMENT* project:

- The MIP/MIPVU approaches are impractical for large corpora, given the analytic detail required to determine the metaphoricity of each and every word in a corpus. One partial solution to this problem is to narrow down the volume of text under analysis, as done for corpus analysis projects by Ogarkova & Soriano (2014) and Semino et al. (2018).
- The analyst's subjectivity affects how the metaphoricity of an expression is determined. This problem is alleviated in the MIP/MIPVU

approaches by relying on the basic (as opposed to literal) meaning of a lexical unit and by making use of external expert knowledge (dictionaries). It is worth noting, however, that advancing the notion of basic instead of literal meaning favours metaphorical readings.

- No precise method for determining conceptual domains is yet available. Indetermination may stem from the need to choose a suitable frame and to choose elements (concepts) within that frame and also the level of generality of the comparison. This multiplicity of possible categories makes it difficult to achieve a reasonable level of inter-annotator agreement.

3 A methodology for detecting and formulating conceptual metaphors

To provide MOMENT project analysts with a method that mitigates the problems outlined above, we established a procedure based on the following principles:

- Working hypothesis formulation and verification at the ME detection stage.
- Partial use of standard methods for metaphorical focus identification.
- Use of external expert knowledge in the form of more extensive use of dictionaries and the additional use of metaphor compendia.
- Implementation of strategies for conceptual metaphor formulation, including domain formulation at two levels of generalization.

Below, in section 3.1, we present the rationale for these four principles and then, in 3.2, we describe their practical application via the annotation method.

3.1 Methodological principles

In this section we present in more detail the four methodological principles that form the basis of our new approach for detecting and formulating conceptual metaphors in corpora.

3.1.1 Using working hypotheses

The problem of time outlined above could be solved by reducing the volume of text to analyse manually, as done in previous corpus analysis projects. However, we did not want to rely on concordance extraction in the first phase of the project as, given the diversity of texts and subjects to be analysed (several types of people diagnosed with mental disorders and several types of mental health professionals), we did not want to be biased by keyword pre-selection.

Therefore, the following strategy was implemented: the analyst is instructed to intuitively pre-select clauses that seem to include MUWs. Their selection as hypothetical MEs can be considered a “working hypothesis”, which is later confirmed or rejected systematically by applying MIP to the words in the clause. The use of working hypotheses is a well-established practice in quantitative and qualitative research, which we apply at a micro-level. Since only the words of hypothetical MEs are analysed – and not every word as in the standard MIP approach – the time needed to apply the MIP is substantially reduced. There is obviously a risk that MEs may go undetected. However, the aim is not to detect each and every ME in the corpus but to identify predominant CMs. The possible loss of some MEs is offset by significant time savings, to the point of ensuring the actual feasibility of the project.

Lastly, one of the specifications to analysts is that a ME should be selected as hypothetical only if it applies to severe mental disorders in accordance with a list of corresponding thematic fields. This specification makes it possible to dismiss a large number of CMs that Geeraerts (2010) would deem problematic as being very conventional or possibly dead, as illustrated in (5a).

3.1.2 Partial application of the MIP

Despite the problems implied by decisions about metaphoricality (see § 2.2.1), like Ogarkova & Soriano (2014) and Semino et al. (2018) we opted to use the MIP, as its advantages outweigh its disadvantages. Even though it favours metaphorical readings and so incurs the risk of overanalysis (because it relies on basic rather than literal meanings), it is well established and provides clear instructions for carrying out analyses at the lexical level.

However, our use of the MIP is partial in that it is not applied to the whole text but only to clauses selected as hypothetical MEs. It is also partial in that we exclude the historicity criterion, as per Steen et al. (2010b).

3.1.3 Using compendia

In MOMENT, the analysts use existing compendia of metaphors or compendia specifically compiled for the project. The compendia are used in the first instance to formulate the CM, and only when the ME does not fit any of the available models are the domains inferred by the analysts.

While the risk exists (as mentioned above) that prior knowledge may bias the analysis, this is always the case when resorting to external information – as done in the MIP when dictionaries are used to determine word meanings. Both dictionaries and compendia are used because they are regarded as expert knowledge resources (the information they contain is viewed more as genuinely useful than as risky). All things considered, the risk of bias is outweighed by the benefits: inter-annotator agreement is facilitated thanks to the mitigation of some of the indeterminacy and subjectivity factors (as discussed in § 2.2.2), results consistent with previous research findings are obtained and the annotator’s task is facilitated and abbreviated by the availability of expert knowledge.

3.1.4 Strategies for conceptual metaphor formulation

We have developed strategies to formulate CMs (i. e., inferring SDs and TDs) when the ME does not seem to fit any of the CMs in the compendia. These strategies are explained to the analysts in the annotation guidelines. We regarded this as necessary for two reasons: firstly, as reported above (§ 2.2.2), as far as we know no instructions have yet been developed for this task, outside of Steen’s suggestion to base inferencing on “prototypical or default knowledge about the source domain” (1999: 71); secondly, better inter-annotator agreement is more likely when a common approach is used for decision-making. Thus, one of the main innovations of our method is the description of systematic strategies for conceptual metaphor formulation to maximize agreement between analysts.

Our strategies are based on substituting words in the ME with other words affording a literal reading. When a verb is used figuratively, we substitute one of its arguments with a more prototypical word (as suggested by Steen 1999) in order to reach a literal reading. In other cases the MUW is substituted by a key concept extracted from the dictionary definition of the word. This is in line with the Pragglejaz Group (2007: 34), which suggests that “Metaphor scholars [...] may profitably use the MIP, especially the step associated with

defining basic meaning, to identify the source and target domains underlying metaphorical words in context”.

Whenever possible, CMs are formulated at two levels of generalization, in line with Ogarkova & Soriano (2014). This approach will be helpful at a later interpretative stage to distinguish between a specific level closer to the field under analysis (in our case, severe mental disorders) and an abstract level more useful for drawing generalizations about broad types of metaphors used in the texts. Moreover, as pointed out by Grady et al. (1999, as cited in Semino et al. 2004: 1291) “conceptual domains are often too general as units of analysis for conceptual metaphors and [...] many mappings are better described as associations between source and target concepts, belonging to distinct domains”.

Our four strategies for CM domain inferencing are described in what follows.

A. Substitution by a prototype argument of a metaphorically used verb

When the metaphorical focus is a verb and its selectional preferences appear to be violated, the analyst determines the prototypical argument(s) for the verb and checks whether substitution of (any of) the argument(s) in the text by the prototype will make the semantic incongruence disappear. If yes, the prototypical argument is established as the SD and the word from the text is established as the TD. As an illustrative example, in applying the MIP to (9a), with *emancipado* (‘emancipated’) detected as the metaphorical focus of the ME, the reasoning is as follows. Using *dolor* (‘pain’) as a prepositional phrase/genitive argument of *emancipar* would violate the verb’s selectional preferences. Since a suitable prototypical argument, inferred from the dictionary definition of *emancipar* (9b), is *autoridad* (‘authority’), using *autoridad* rather than *dolor* in the text would meet the verb’s selectional preferences. Consequently, the CM is formulated using *dolor* as the TD and *autoridad* as the SD (9c).⁴

⁴ In the conceptual metaphor theory literature, concepts and domains are conventionally represented in uppercase. The same convention is used by us, but in example’s explanations, for stylistic reasons, we use lowercase for the word which evokes the concept.

- (9) a. *no hay nadie **emancipado** del dolor.*
 ‘nobody is **emancipated** from pain.’
- b. Emancipar 1: Liberar de la autoridad legal paterna, de la tutela, de la servidumbre o de otro tipo de subordinación o dependencia
 ‘Emancipate 1: To grant freedom from parental legal authority, guardianship, servitude or another type of subordination or dependence’.
- c. EL DOLOR ES UNA AUTORIDAD
 PAIN IS AN AUTHORITY

The wording of the argument will preferably be extracted from the dictionary definition of the focus (for example *autoridad* from the dictionary definition of *emancipar*). Note that, in this strategy, while domains may be generated from the focus, the focus itself is not established as a domain of the underlying metaphor (e. g. *emancipar* itself is not considered a source or a target domain).

B. Substitution by keywords from the dictionary definition of the focus in a lexicalized metaphor

When the metaphorical focus is not a verb, or when it is a verb but strategy A is not applicable, the analyst formulates an operational comparison between the contextual and basic meanings of the focus. If the contextual meaning of the focus is lexicalized (i. e., it corresponds to one of the meanings of the focus in the reference dictionary), the metaphor formulation is drawn from the contextual and basic dictionary definitions. A word representing the contextual meaning is annotated as the TD, and a word corresponding to the basic meaning is annotated as the SD.

In (10a), *estigma* is detected as the focus on applying the MIP. The contextual meaning corresponds to word sense 2 (10b), whereas the basic meaning (10c) is more concrete, more visible and related to the body. Therefore, the CM is formulated using *deshonra* (‘dishonour’) as the TD and *marca en el cuerpo* (‘mark on the body’) as the SD (10d), both chosen as the more explanatory lexical units in the dictionary definitions. As in strategy A before, the focus itself (*estigma*) is not established as a domain of the underlying metaphor.

- (10) a. *Hace más de un año que apareció el **estigma**.*
 ‘The **stigma** appeared more than a year ago.’
- b. Estigma 2: Motivo de deshonra o de mala fama.
 ‘Stigma 2: Reason for dishonour or bad reputation.’

- c. Estigma 1: Marca o señal en el cuerpo.
'Stigma 1: Mark or scar on the body.'
- d. LA DESHONRA ES UNA MARCA EN EL CUERPO
DISHONOUR IS A MARK ON THE BODY

C. Substitution by keywords from the dictionary definition of the focus in a non-lexicalized metaphor

If the contextual meaning is not lexicalized (i. e., it does not correspond to any of the meanings of the focus in the reference dictionary), the analyst will infer meaning from the context or co-text. The word representing the contextual meaning of the focus is then annotated as the TD, and the focus itself or a label corresponding to its basic meaning is annotated as the SD.

On applying the MIP in (11a), *pegatina*, an informal synonym for *adhesivo* ('sticker'), is detected as the focus. From prior context it is inferred that this word refers metaphorically (and pejoratively) to a person's diagnosis (*diagnóstico* in Spanish), whereas its only meaning in the dictionary is (11b). The CM is thus formulated using *diagnóstico* as the TD and *adhesivo* as the SD (11c), since the more formal label extracted from the dictionary definition is preferred.

- (11) a. *Tomé la decisión de quitarme la **pegatina**.*
'I made the decision to get rid of my **sticker**.'
- b. *Pegatina 1: Adhesivo pequeño que lleva impreso un texto o una imagen.*
'Sticker 1: Small adhesive piece of paper with text or pictures printed on it.'
- c. UN DIAGNÓSTICO ES UN ADHESIVO
A DIAGNOSIS IS A STICKER

D. Substitution and adscription to a more general mapping

The strategies presented above generate metaphorical comparisons at the specific level. But in some cases the comparisons may be more general, leading to the formulation of appropriate general-level metaphors. In other cases the comparison is recognized as an ontological correspondence for a CM (Lakoff & Johnson 1980), inasmuch as one can think of other mappings between concepts in the domains. These two cases of generalization correspond to the two cases of specification of general metaphors posited by Ogarkova & Soriano (2014), which they call "special case" and "entailment",

respectively. Here we present an example of the latter.

In (12a), *apostamos* is one of the expression's MUWs. Following the above-described strategy B, the CM is initially formulated as (12b). However, the basic meaning of *apostar* (12c) in the reference dictionary indicates that the concept belongs to the gambling domain. Several elements can be distinguished within this complex frame: gambling, gaming houses, money, getting rich, going bankrupt, etc. Therefore, the mapping in (12b) can be reasonably regarded as one of several possible mappings in a more general analogy between gambling and life.⁵ Lakoff & Johnson (1980: 51), in fact, postulate LIFE IS A GAMBLING GAME as a conventional metaphor that covers expressions such as *I'll take my chances, the odds are against me, I've got an ace up my sleeve*, etc. Since equivalents for these and similar expressions exist in Spanish, it seems reasonable to formulate a more general metaphor as in (12d). The fact that the CM underlying *apostamos* is richly annotated at two levels – specific in (12b) and generic in (12d) – captures both aspects of the conceptual mapping.

- (12) a. *Elegimos un camino peligroso, **apostamos** fuerte, alto...*
 ‘We choose a dangerous road, we **bet** hard, high...’
- b. ARRIESGAR ES APOSTAR
 TO RISK IS TO BET
- c. Apostar 2: Referido a una cantidad de dinero, arriesgarla para poder participar en el juego que consiste en acertar el resultado de algo, de forma que, si se acierta, se recibe una cantidad de dinero mucho mayor. ‘Bet 2: Referring to a quantity of money, risking it to participate in the game that consists of getting the result of something right, so that, if right, a much larger amount of money is received.’
- d. LA VIDA ES UN JUEGO DE AZAR
 LIFE IS A GAMBLING GAME

3.2 Annotation method

The method for manually annotating figurative expressions in the MOMENT project is organized along mutually exclusive paths, consisting of steps with explicit instructions for detecting and formulating three possible conceptual figures: CMs, conceptual metonymies and metaphorical similes. Although in this paper we focus on CMs, below we also briefly describe the analytical steps

⁵ The concept LIFE IS chosen as TD because the co-text indicates that the speaker is talking about his life trajectory.

concerning conceptual metonymies and metaphorical similes. The overall structure of the process is depicted in Figure 1.

Annotators are issued a document of annotation guidelines and several complementary documents. The former details the procedure described below and the latter are the following:

1. List and description of thematic fields specific to severe mental disorders
2. List of lexical units excluded as potential MUWs
3. List of comparison markers for simile detection
4. Compendium of mental health CMs
5. Compendium of mental health conceptual metonymies
6. Compendium of general purpose CMs
7. Compendium of general purpose conceptual metonymies.

Documents 1–5 were specifically compiled for the project. Document 1 has been drawn up based on the authors' previous work on schizophrenia metaphors (Climent & Coll-Florit 2017; Coll-Florit et al. 2019); it lists and briefly describes semantic fields considered relevant for the analysis of discourse in mental disorders (the life of the affected people, the symptoms, the related emotions, the medication and professional intervention, social prejudices and discrimination, etc). Document 2 lists certain word classes, delexicalized words and data-specific terms to be excluded from the analysis, as per Semino et al. (2018: 59).⁶ Document 3 is based on the Real Academia Española (2009: 3408–3420) grammar of the Spanish language. Documents 4 and 5, compiled from Barcelona (1986), Semino (2008: 178–190), Climent & Coll-Florit (2017) and Coll-Florit et al. (2019), consist of metaphor formulations with representative examples. Document 6 is the Master Metaphor List (Lakoff et al. 1991). Document 7 was extracted

⁶ These are the following: (1) terms very commonly used in the field such as *brote* ('flare-up', literally 'sprout'), which were metaphorical in origin but have become medical terminology; (2) prepositions when carrying no semantic content, with some exceptions, such as the locative uses of *en* ('in'); (3) lexical verbs functioning as auxiliary or modal markers such as *tener* ('have') or *acabar* (literally 'bring to an end'); and (4) interjections, e. g. *qué diablos* ('what the hell'), as they can be analysed metaphorically but have become extremely conventional.

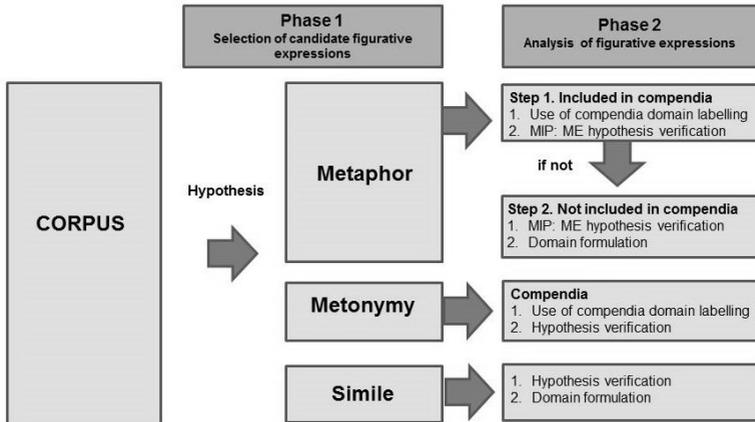


Figure 1. Overall structure of the annotation method

from Littlemore & Tagg (2016) (who drew, in turn, on Radden & Kövecses 1999).

The analysts use two Spanish language dictionaries, CLAVE (Maldonado 2012) and *Diccionario de la Lengua Española* [DLE] (Real Academia Española 2001), chosen for the following reasons: both are online so lookup is facilitated; CLAVE is a recent usage dictionary with clear-cut word meanings and illustrative examples; and DLE is the standard Spanish language reference dictionary. Instructions are to preferably use CLAVE and to resort to DLE in cases of doubt.

The method consists of two main phases: selecting the hypothetical figurative expression and analysing the hypothetical figurative expression. Sentences are randomly presented to the analyst. While the immediate context is undeniably primordial, each sentence is presented with the immediately preceding and subsequent sentences, and analysts are also instructed to consult the original text (easily accessed through hypertext links) if doubts remain.

3.2.1 Phase 1: Selection of candidate figurative expressions

The analyst reads the sentence and its context to capture the general meaning and, checking against the list of thematic fields for severe mental disorders, decides whether the sentence contains one or more candidate figurative expressions (those perceived not to belong are rejected). The following general clues are used to hypothesize the occurrence of figurative expressions and to distinguish between metaphors, metonymies and similes:

- **CMs.** An indirect or non-literal use is made of a word or a group of words that, in the context, seems to express some kind of comparison or resemblance between concepts in such a way as to make the discourse more expressive or understandable.
- **Conceptual metonymies.** A noun or a noun phrase represents and/or replaces another noun or noun phrase belonging to the same domain of knowledge and the concepts represented are related by a spatial, temporal, causal or part-whole contiguity.
- **Metaphorical similes.** The sentence contains a comparison marker.

Hypothetical figurative expressions are extracted as clauses from the sentence with enough immediate context to be understandable.

3.2.2 Phase 2: Analysis of figurative expressions

The analysis evolves along separate no-return paths referring to the three kinds of conceptual figures postulated by the analysts on the basis of hypotheses.

Conceptual metaphor analysis

CMs are annotated in one of two ways, depending on whether the metaphor is included or not included in compendia (Step 1 or Steps 1 and 2, respectively).

Step 1: Metaphors included in compendia

In this step, in the first place the CM domains are provisionally annotated and, in the second place, expressions are analysed in order to verify or reject both their status as an ME and their correspondence to the tentative CM. The process is as follows: for each hypothetical ME, the analyst checks the

metaphor compendia to determine if a corresponding suitable CM exists. Analysts are instructed to first check the compendium of mental health metaphors and only use the general purpose compendium if the metaphor is not found in the former. If a suitable CM is found in the compendia, this tentative CM is used (i. e. domains are annotated) at the generic level; otherwise the analyst proceeds to Step 2. The specific-level formulation is determined from the text – in a straightforward way if both domains are explicit or applying one of the domain inference strategies (as described in § 3.1.4 above). Last but not least, the metaphoricity of the hypothetical ME is verified or rejected by application of the MIP to all its lexical units. If at least one of these units is judged to be an MUW, the metaphoricity of the ME is verified; and if the ME is judged to be congruent with the tentatively formulated CM, then the annotation is deemed valid. Otherwise, the provisional annotations are cancelled, and the process goes to Step 2.

As can be seen, this step proceeds in reverse order to what could be expected, since determining if an expression is metaphorical typically comes before trying to label the corresponding CM. Our experience in metaphor annotation has shown that dealing first with intuitions based on previous work (compendia) and then verifying or counter-verifying them results in more consistent labellings and annotation time saving. This occurs mainly because the same generic CMs included in compendia are being repeated in texts: e. g. *THE DISORDER IS A LOCATION* is instantiated in several specific locations: hell, pit, hole, gutter, labyrinth, etc. Therefore, what differentiates this approach from past practices of CM accounting (e. g. Steen 1999; 2007) is that our method is based on metaphor compendia in the first instance.

As an example of Step 1, in (13a) the analyst selects the clause *conocer el infierno* (13a) as an hypothetical ME, with *infierno* as the potential focus. The mental health metaphor compendium includes *EL TRASTORNO ES UN LUGAR* ('*THE DISORDER IS A LOCATION*'), considered applicable to the expression. Thus, at the generic level, *trastorno* is annotated as the TD and *lugar* as the SD. As for the specific level, in coherence with the text, *infierno* ('HELL') is chosen as the SD and the TD results from hypothesis verification. The analyst judges that the contrast between meanings 8 (13b) and 1 (13c) for *infierno* in the DLE dictionary is metaphorical in accordance with MIP specifications, and that this contrast is congruent with the provisionally postulated metaphor, and so the hypothesis is verified. Moreover, applying the second domain inference strategy, i. e., substitution based on keywords from the dictionary definition of the focus in a lexicalized metaphor (§ 3.1.4), the specific TD is

inferred by comparing the two dictionary definitions. From (13b) *sufrimiento* is considered to be the more faithful representation of the meaning conveyed by the text and so the CM underlying (13a) is formulated as (13d).

- (13)
- a. *Lo que no sabía [...] es que a los pocos días iba a conocer el **infierno***
‘I didn’t know that in a few days I was going to be in **hell**’
 - b. Infierno 8: Lugar o situación que causa gran sufrimiento o malestar.
‘Hell 8: A place or situation causing extreme suffering or distress’.
 - c. Infierno 1: En la doctrina tradicional cristiana, lugar donde los condenados sufren, después de la muerte, castigo eterno.
‘Hell 1: In traditional Christian doctrine, the place where the damned suffer eternal punishment after death.’
 - d. Specific level: EL SUFRIMIENTO ES EL INFIERNO (SUFFERING IS HELL).
Generic level: EL TRASTORNO ES UN LUGAR (THE DISORDER IS A LOCATION).

Step 2: Metaphors not included in compendia

When a compendium metaphor is not identified, the procedure is reversed, i. e., the focus of the ME is first detected and the CM domains are then inferred. The analyst first applies the MIP to all the words in the hypothetical ME. If a word is eligible, it is annotated as the focus and the hypothesis is verified; the analyst then formulates the domains of the underlying CM. If no MUWs are detected by the MIP, the hypothesis is rejected and the process ends. Conceptual domain labelling consists of an initial formulation at the specific level – by application of the appropriate domain inference strategy (see § 3.1.4) – followed by a generic level annotation according to one of the following possibilities: the specific metaphor is either a subtype of a more abstract metaphor (as in 13) or is an ontological correspondence of a CM (as in 12). If a generic-level labelling is not found to be possible, the ME remains just annotated at the specific level. This is the case with examples (10) and (11), which can be regarded as *one-shot metaphors* (Steen 1999: 58–59).

Conceptual metonymy analysis

Analogously to Step 1 for CMs, the analysis consists of first determining and then verifying the metonymy. In line with Brdar (2018), who states that there are no new metonymies only new instances of known types created by analogy, it is assumed that the general compendium of conceptual metonymies

provides comprehensive coverage of all possible types at the generic level. The analyst determines whether the source term (in the text) and the target concept instantiate a compendium metonymy. If so, the source term is annotated as the focus, the compendium metonymy is provisionally annotated at the generic level and the concepts or domains directly emerging from the text are annotated at the specific level. Otherwise, the hypothesis is rejected and the process ends. The hypothesis is verified by analysing the metonymic focus using a method adapted from (Biernacka 2013: 117, cf. Littlemore & Tagg 2016). The procedure is similar to the MIP, except that, in this case, concepts are related by spatial, temporal, causal or part-whole contiguity.

By way of an example, in (14), *mundo* is postulated as the focus and is determined, in this context, to instantiate (14b), a metonymy included in the general compendium. Consequently, at the generic level *lugar* is annotated as the SD and *habitantes de un lugar* as the TD. To determine the specific level, the dictionary meanings of *mundo* are checked. Both relevant meanings are lexicalized, (14c) as the basic concept and SD and (14d) as the TD. From the definitions, the metonymy is formulated at the specific level as in (14e).

- (14) a. *Es el mundo, que no está preparado.*
 ‘It is the **world**, which is not ready yet’.
- b. EL LUGAR POR LOS HABITANTES DE UN LUGAR
 ‘THE PLACE FOR INHABITANTS’
- c. Mundo 3: Planeta o astro, esp. referido a la Tierra.
 ‘World 3: Planet or heavenly body, esp. the Earth.’
- d. Mundo 4: Conjunto o sociedad de los seres humanos.
 ‘World 4: All people or society’.
- e. EL PLANETA TIERRA POR LA SOCIEDAD
 PLANET EARTH FOR SOCIETY

Metaphorical simile analysis

Following Semino et al. (2018: 282), we understand similes as “the explicit linguistic formulation of a comparison between two unlike entities, usually signalled by expressions such as ‘like’ or ‘as if’, as in ‘Cancer is like a journey’”. In similes there is no contrast between basic and contextual meanings of the source, therefore it differs from CMs in that in this case language is used “directly” (Steen et al. 2010b: 11). Our annotation guidelines consider a simile to be metaphorical if it establishes a comparison between a target abstract term and a source concrete term, while it is considered

non-metaphorical if it just establishes some kind of similitude between entities.

Taking a comparison connector as indicating a hypothetical conceptual simile, the analyst first determines whether the contrast between the two terms is metaphorical. If so, the hypothesis is verified and the first term of the comparison is provisionally annotated as the TD and the second term as the SD. If the comparison between the two terms is literal, then the process ends. Finally, if applicable, the generic level is established as in CMs.

For instance, in (15) the presence of a simile is indicated by *como* ('like'), which reflects a comparison of the concepts denoted by *enfermo mental* and *reloj*. The analyst judges the contrast to be metaphorical, given that it is asymmetrical, highlights a single property and contains an explanation or elaboration. The concept of *la persona diagnosticada*⁷ is annotated as the TD and *reloj* is annotated as the SD (15b). Finally, the conventional metaphor (15c) is formulated at the generic level, on the basis that the comparison in the text is judged to be an instance of it.

- (15) a. *El enfermo mental [...] como un reloj debe marcar las horas.*
 'The mental patient [...] like a clock has to mark time'.
 b. LA PERSONA DIAGNOSTICADA ES UN RELOJ
 'THE DIAGNOSED PERSON IS A CLOCK'
 c. UNA PERSONA ES UNA MÁQUINA
 'A PERSON IS A MACHINE'

4 Reliability testing

To assess the validity of the method and to empirically assess its reliability, inter-annotator agreement was tested. Details of the test corpus, the analysts and the coding procedure are described below, followed by the results, reflecting the degree of statistical confidence.

4.1 Test corpus

The research was conducted in the framework of a project whose primary focus, as mentioned earlier, is the analysis of mental health metaphors. The corpus used for inter-annotator agreement testing included Spanish texts (4,143 words in 200 sentences) produced by two subjects with a severe mental

⁷ In the MOMENT project we use this term to refer to the people with a mental health condition instead of *enfermo mental* or similar derogatory or disrespectful words.

disorder, published in a book describing personal stories and experiences of the mental disorders (*El libro de Radio Nikosia*; Vv. Aa 2005).

4.2 Analysts

Corpus annotation was independently performed by the authors of the annotation method described in this paper, both linguists and both experts in conceptual metaphor theory, who have worked together in the area of mental health metaphors for four years. Before testing, the analysts, thoroughly versed in the annotation procedure, had discussed the method in several meetings.

4.3 Annotation interface and analysed items

Metaphors were coded in ten interface fields in Excel files (Table 1).

The reliability test examined inter-annotator agreement for the following main items:

- a) Identification of metaphorical language use. For the 200 sentences in the test corpus, we calculated agreement between the analysts regarding whether a sentence included metaphorical language reflecting the severe mental disorders field.
- b) Use of metaphor compendia. To assess the usefulness of the metaphor compendia as expert knowledge, we calculated agreement between the analysts regarding sentences including metaphors included in compendia.
- c) MUW annotation. For sentences that both analysts hypothesized as metaphorical, we calculated their agreement in coding the MUWs. It is important to note that this is the only item evaluated in reliability testing for MIP (Pragglejaz Group 2007) and MIPVU (Steen et al. 2010b).
- d) Conceptual domain formulation. For the identified metaphors, we calculated agreement between the analysts in the formulation of conceptual domains (SD and TD). This item has not been previously evaluated in reliability tests of metaphor annotation. Although Ogarkova & Soriano (2014) report agreement results on metaphor labelling, they are actually based on a closed inventory of CMs into which to try and classify MEs. Therefore, technically speaking, it can be regarded as metaphor classification, inasmuch as there is no application of strategies for domain formulation.

Table 1. Annotation interface for the inter-annotator agreement test

Field	Value	Description
ID-String	Natural number (1)	Identifier for each sentence in the corpus.
ID-Line	Rational number (1.1)	Identifier for each line in the Excel file. A sentence is split in several lines if it contains more than one ME.
ID-Author	Natural number (1)	Identifier for the author of the text.
String	Text	Sentence. It can be repeated as many times as a ME is detected.
Metaphorical expression (ME)	Text / NO	Clause for which metaphor use is hypothesized (in the severe mental disorder field). Standing alone it must be meaningful, although it may be discontinuous in relation to the original text. If there is no metaphorical language, this field is coded as NO.
Figure	MF, MN, SI	Figure descriptor. MF: metaphor MN: metonymy SI: simile
Focus	Text	MUWs in the ME.
Conceptual domains	Text	Formalization of the conceptual domains for the metaphor: specific and generic.
Compendia	C, N	Origin of the figure. C: included in metaphor compendia. N: not included in metaphor compendia.
Comments	Text	Open field.

Table 2. Inter-annotator agreement and statistical confidence scores

Items	Agreement	Cohen's Kappa
Identification of metaphorical language use referring to severe mental disorders	91%	0.81
Identification of metaphors included in compendia	87%	0.70
MUW annotation	97.6%	0.79

4.4 Reliability results

In this section we present the results of the agreement test between analysts before discussion. To investigate possible bias in the individual analysts, we computed the kappa test statistic, the reliability measure most widely used in previous studies of metaphorical language annotation (Markert & Nissim 2003; Pragglejazz Group 2007; Steen et al. 2010b). More specifically, we used Cohen's kappa, given that we were testing agreement between a pair of analysts. In interpreting kappa values, the literature indicates that values in the range 0.61–0.80 point to substantial agreement.

Our test results (Table 2) point to a high degree of agreement between the analysts, at 91% for the identification of metaphorical language use related to the semantic fields of severe mental disorders (kappa value 0.81), and at 87% for the identification of metaphors included in compendia (kappa value 0.70). Regarding the annotation of MUWs, the percentage of agreement rises to 97.6% (kappa value 0.79). Taking kappa values reported for previous MIP and MIPVU inter-annotator agreement tests as the gold standard in MUW annotation, the Pragglejazz Group (2007) reported kappa values of 0.62 for conversation texts and of 0.72 for news texts (MIP), while Steen et al. (2010b) reported kappa values ranging from 0.70 to 0.96 for different text genres (academic texts, fiction, news and conversation), with the conversation genres obtaining the lowest values (MIPVU). Our kappa value results, therefore, are better than those for MIP and similar to those for MIPVU.

Finally, Table 3 shows the level of agreement in the formulation of conceptual domains, a concept not previously evaluated. For this analysis, we considered fragments for which both analysts identified metaphorical language use. Since conceptual domains are not a closed tagset, the categories are not mutually exclusive as required for kappa calculations and,

Table 3. Inter-annotator agreement in conceptual domain formulation

Both SD and TD	Agreement		
	At least one domain	SD	TD
71%	87%	78%	80%

consequently, we only report the percentage agreement between analysts. Note that two formulations for the same domain were considered equivalent if there was agreement at either the specific or the generic level. Additionally, two formulations were recognized as equivalent if they used synonyms or different grammatical categories but clearly referred to the same entity or concept (e. g. in Spanish *transferencia* (n) / *transferir* (v) [‘transfer’]).

As can be observed in Table 3, the analysts proposed the same formulation for both metaphor domains in 71% of the cases and agreed regarding at least one domain in 87% of the cases. Interestingly, there was no substantial difference regarding agreement as to the SD (78%) and the TD (80%).

Finally, it is worth mentioning that, when the annotation includes domain formulation at two levels of generalization, there is agreement in both levels in 76% of the cases. Furthermore, when there is agreement in only one level (in the 24% remaining cases), we did not find significant differences in agreement between the specific and the generic levels.

5 Reproducibility of the method

The new methodological approach presented in this paper offers systematicity for both ME detection and CM formulation, which favours its replicability to other corpora. However, there are still some variables that could introduce noise in the reproducibility of the method.

Regarding ME detection, different analysts may choose different basic and contextual dictionary definitions. In a related way, some definitions are ambiguous between a basic and a figurative meaning. Nevertheless, note that these are not new limitations but inherent to the application of the MIP. With reference to CM formulation, different researchers may choose different keywords from dictionary definitions when applying strategies for domain formulation. Therefore, some degree of subjectivity can still be expected in applying the method, although this is a limitation which is unavoidable when dealing with manual annotation of corpora.

6 Conclusions

In the context of research conducted in the framework of MOMENT, a project that aims to analyse a large corpus of mental health texts in the light of conceptual metaphor theory, we have developed a feasible and reliable methodology for detecting MEs in discourse and establishing the underlying cross-domain mappings.

In designing our method, we fundamentally pursued two aims: (i) to maximize agreement between analysts by reducing the impact of subjectivity and indetermination as much as possible, given that conceptual domain inference is a highly elusive task for which standard methods offer little guidance; and (ii) to render the task feasible in terms of time. Our method is structured as a workflow, with mutually exclusive paths, designed on the basis of the following principles: use of working hypotheses, partial use of the MIP (Pragglejaz Group 2007) to establish metaphoricity, use of external expert knowledge (dictionaries and compendia) and the implementation of specific strategies to infer underlying conceptual domains.

Thus, this method presents some innovations and makes new contributions both in metaphor detection and formulation. In regard to detection, the use of working hypotheses at the ME stage substantially reduces the time needed to apply the MIP, so that it does not have to be applied to each and every word in a corpus. In relation to formulation, the use of metaphor compendia in the first instance facilitates and abbreviates the annotator's task and fosters inter-annotator agreement. Moreover, and crucially, this method presents several systematic strategies for conceptual metaphor formulation, based on a more intensive use of dictionaries. This last point is especially relevant taking into account that the field of metaphor studies lacks a standardized method for conceptual metaphor formulation.

Satisfactory results were obtained when we tested our method for inter-annotator agreement with a 4,143-word test corpus. Regarding ME detection, the statistical confidence values were better than those reported for the MIP (Pragglejaz Group 2007) and similar to those reported for the derivative MIPVU (Steen et al. 2010b). With respect to domain formulation, agreement was reached in around three quarters of the cases, a satisfactory result considering that conceptual domains are not a closed tagset but are open to the annotator's subjectivity. Therefore, we have assessed the validity of this new methodology that, although applied to a corpus of mental health texts in this paper, can be easily applied to essentially any corpus.

Acknowledgements

This research was conducted in the framework of the project *MOMENT: Metaphors of severe mental disorders. Discourse analysis of affected people and mental health professionals*, funded by the Spanish National Research Agency (Agencia Estatal de Investigación, AEI) and the European Regional Development Fund (ERDF), within the National Programme for Research Aimed at the Challenges of Society. Ref. FFI2017-86969-R (AEI/ERDF, EU).

We would like to thank the two anonymous reviewers whose comments and suggestions greatly helped to improve this manuscript. All remaining flaws are ours.

Abbreviations

CM	conceptual metaphor
ME	metaphorical expression
MIP	metaphor identification procedure
MIPVU	metaphor identification procedure Vrije University
MUW	metaphorically used words
SD	source domain
TD	target domain

References

- Barcelona, Antonio. 1986. On the concept of depression in American English: A cognitive approach. *Revista Canaria de Estudios Ingleses* 12. 7–33.
- Biernacka, Ewa. 2013. *The role of metonymy in political discourse*. Milton Keynes: The Open University. (Doctoral dissertation).
- Brdar, Mario. 2018. On the novelty of novel metaphors and metonymies. (Plenary lecture at AELCO/SCOLA 2018, 11th International Conference of the Spanish Cognitive Linguistics Association, 19 October 2018).
- Climent, Salvador & Coll-Florit, Marta. 2017. La metáfora conceptual en el discurso psiquiátrico sobre la esquizofrenia [Conceptual metaphor in psychiatric discourse on schizophrenia]. *Ibérica* 34. 187–208. (Special Issue on Metaphor in Science and Technology).

- Coll-Florit, Marta & Climent, Salvador & Correa-Urquiza, Martín & Hernández, Eulàlia & Oliver, Antoni & Pié, Asun. 2018. MOMENT: Metáforas del trastorno mental grave: Análisis del discurso de personas afectadas y profesionales de la salud mental [MOMENT: Metaphors of severe mental disorder: Discourse analysis of affected people and mental health professionals]. *Procesamiento del Lenguaje Natural* 61. 139–142.
- Coll-Florit, Marta & Miranda, Xavier & Climent, Salvador. 2019. Metáforas de la esquizofrenia: Un estudio sobre el discurso de afectados y profesionales [Metaphors of schizophrenia: Discourse analysis of sufferers and professionals]. *Revista Española de Lingüística Aplicada (RESLA) / Spanish Journal of Applied Linguistics (SJAL)* 32(1). 1–31.
- Ding, Yan. 2011. *Metaphors of sadness: Intraconceptual and interconceptual variation*. Hong Kong: University of Hong Kong. (Doctoral dissertation).
- Geeraerts, Dirk. 2010. *Theories of lexical semantics*. Oxford: Oxford University Press.
- Grady, Joseph & Oakley, Todd & Coulson, Seanna. 1999. Blending and metaphor. In Gibbs, Raymond W. Jr. & Steen, Gerard J. (eds.), *Metaphor in cognitive linguistics*, 101–124. Amsterdam: John Benjamins.
- Heywood, John & Semino, Elena & Short, Mick. 2002. Linguistic metaphor identification in two extracts from novels. *Language and Literature* 11(1). 35–54.
- Kövecses, Zoltán. 2002. *Metaphor: A practical introduction*. New York: Oxford University Press.
- Lakoff, George & Espenson, Jane & Schwartz, Alan. 1991. Master Metaphor List. Berkeley, CA: Cognitive Linguistics Group University of California, Berkeley. (Technical report).
- Lakoff, George & Johnson, Mark. 1980. *Metaphors we live by*. Chicago: University of Chicago Press.
- Lakoff, George & Turner, Mark. 1989. *More than cool reason: A field guide to poetic metaphor*. Chicago: University of Chicago Press.
- Littlemore, Jeannette & Tagg, Caroline. 2016. Metonymy and text messaging: A framework for understanding creative types of metonymy. *Applied Linguistics* 39(4). 481–507.
- Maldonado, Concepción (ed.). 2012. *Diccionario CLAVE: Diccionario de uso del español actual* [CLAVE dictionary: Dictionary of Spanish contemporary usage]. Cataño: Ediciones SM.
(<http://clave.smdiccionarios.com/app.php>) (Accessed 2019-03-15).
- Markert, Katja & Nissim, Malvina. 2003. Corpus-based metonymy analysis. *Metaphor and Symbol* 18. 175–188.
- Ogarkova, Anna & Soriano, Cristina. 2014. Variation within universals: The ‘metaphorical profile’ approach to the study of ANGER concepts in English, Russian and Spanish. In Musolff, Andreas & MacArthur, Fiona & Pagani, Giulio (eds.), *Metaphor and intercultural communication*, 93–116. London: Bloomsbury.

- Pragglejaz Group. 2007. MIP: A method for identifying metaphorically used words in discourse. *Metaphor and Symbol* 22(1). 1–39.
- Radden, Günter & Kövecses, Zoltán. 1999. Towards a theory of metonymy. In Panther, Klaus-Uwe & Radden, Günter (eds.), *Metonymy in language and thought*, 17–59. Amsterdam: John Benjamins.
- Real Academia Española. 2001. *Diccionario de la lengua española* [Dictionary of the Spanish language]. (<http://www.rae.es/rae.html>) (Accessed 2019-03-15).
- 2009. *Nueva gramática española: Sintaxis II* [New Spanish grammar: Syntax II]. Madrid: Espasa Libros.
- Rundell, Michael & Fox, Gwyneth (eds.). 2002. *Macmillan English dictionary for advanced learners*. Oxford: Macmillan Education.
- Semino, Elena. 2008. *Metaphor in discourse*. Cambridge: Cambridge University Press.
- Semino, Elena & Demjen, Zsófia & Hardie, Andrew & Payne, Sheyla & Rayson, Paul. 2018. *Metaphor, cancer and the end of life: A corpus-based study*. London: Routledge.
- Semino, Elena & Heywood, John & Short, Mick. 2004. Methodological problems in the analysis of metaphors in corpus conversations about cancer. *Journal of Pragmatics* 36(7). 1271–1294.
- Soriano, Cristina. 2012. La metáfora conceptual [Conceptual metaphor]. In Ibarretxe-Antuñano, Iraide & Valenzuela, Javier (eds.), *Lingüística cognitiva*, 97–120. Barcelona: Anthropos.
- Steen, Gerard J. 1999. From linguistic to conceptual metaphor in five steps. In Gibbs, Raymond W. Jr. & Steen, Gerard J. (eds.), *Metaphor in cognitive linguistics*, 55–77. Amsterdam: John Benjamins.
- 2007. Finding metaphor in discourse: Pragglejaz and beyond. *Cultura, Lenguaje y Representación / Culture, Language and Representation* 5. 9–25.
- Steen, Gerard J. & Biernacka, Ewa & Dorts, Aletta G. & Kaal, Anna & López Rodríguez, Clara I. & Pasma, Trijntje. 2010a. Pragglejaz in practice: Finding metaphorically used words in natural discourse. In Low, Graham & Todd, Zazie & Deignan, Alice & Cameron, Lynne (eds.), *Researching and applying metaphor in the real world*, 165–184. Amsterdam: John Benjamins.
- Steen, Gerard J. & Dorts, Aletta G. & Herrmann, Berenike J. & Kaal, Anna & Krennmayr, Tina & Pasma, Trijntje. 2010b. *A method for linguistic metaphor identification: From MIP to MIPVU*. Amsterdam: John Benjamins.
- Stefanowitsch, Anatol. 2006. Words and their metaphors: A corpus-based approach. In Stefanowitsch, Anatol & Gries, Stefan Th. (eds.), *Corpus-based approaches to metaphor and metonymy*, 61–105. Berlin: Mouton de Gruyter.
- Valenzuela, Javier & Soriano, Cristina. 2005. Cognitive metaphor and empirical methods. *Barcelona English Language and Literature Studies (BELLS)* 14. 1–19.

Vv. Aa. 2005. *El libro de Radio Nikosia: Voces que hablan desde la locura* [The Radio Nikosia book: Voices speaking from madness]. Barcelona: Gedisa.

Contact information:

Marta Coll-Florit
Salvador Climent
Faculty of Arts and Humanities
Universitat Oberta de Catalunya
Av. Tibidabo 47
08035 Barcelona
Catalonia, Spain
e-mail: mcollfl(at)uoc(dot)edu
 scliment(at)uoc(dot)edu