Meanings are like Onions: a Layered Approach to Metaphor Processing⋆

Silvia Cappa[1], Anna Sofia Lippolis[1,2] and Stefano Zoia[3]

[1]Institute for Cognitive Sciences and Technologies (ISTC), CNR, Rome, Italy
[2]University of Bologna, Italy
[3]University of Turin, Italy

Abstract

Metaphorical meaning is not a flat mapping between concepts, but a complex cognitive phenomenon that integrates multiple levels of interpretation. In this paper, we propose a stratified model of metaphor processing that treats meaning as an onion: a multi-layered structure comprising (1) contextual information, (2) conceptual blending analysis, and (3) pragmatic analysis. This three-dimensional framework allows for a richer and more cognitively grounded approach to metaphor interpretation in computational systems. At the first level, metaphors are annotated through contextual metadata. At the second level, we model conceptual combinations, linking components to emergent meanings. Finally, at the third level, we introduce a pragmatic vocabulary to capture speaker intent, communicative function, and contextual effects, aligning metaphor understanding with pragmatic theories. By unifying these layers into a single formal framework, our model lays the groundwork for computational methods capable of representing metaphorical meaning beyond surface associations—toward deeper, more context-sensitive reasoning.

Keywords

Metaphors, Metaphor representation, Pragmatics, AI

1. Introduction

Metaphors pervade human communication and cognition, extending far beyond mere linguistic decoration. As cognitive tools, they grant privileged access to implicit knowledge structures that might otherwise remain hidden [1]. By mapping relationships between concepts, metaphors serve as bridges that both reveal and reshape our conceptual frameworks—think of how we say that we *spend*, *save*, or *waste* time, implicitly assuming it is a finite resource.

Despite rapid progress in natural language processing, computational metaphor analysis continues to face five intertwined challenges rooted in the very knowledge structures metaphors invoke:

1. *Data scarcity and representational gaps*. Datasets accounting for many metaphorical phenomena are scarce, and building new ones is (i) resource-intensive, and (ii) hindered by frameworks that go no further than simple domain mappings.
2. *Contextual insensitivity*. While the Conceptual Metaphor Theory (CMT) developed by Lakoff and Johnson [2] dominates in computational accounts of metaphors, it often fails to capture, among other things, how context shifts a metaphor's meaning in discourse.
3. *Evaluation and standardization*. Definitions, metrics, and supported linguistic forms (nominal, verbal, adjectival) in metaphor processing vary wildly across studies [3].
4. *Theoretical fragmentation*. Competing accounts (e.g., CMT vs. interactional or embodiment theories) illuminate different aspects of metaphorical phenomena but rarely integrate pragmatics—what a metaphor does in conversation—often goes unmodeled despite Speech Act Theory (SAT) being a long-standing account of these processes and widely adopted in the literature [4].

Proceedings of the Joint Ontology Workshops (JOWO) - Episode XI: The Sicilian Summer under the Etna, co-located with the 15th International Conference on Formal Ontology in Information Systems (FOIS 2025), September 8-9, 2025, Catania, Italy

†Equal contribution.

silviacappa@cnr.it (S. Cappa); annasofia.lippolis2@unibo.it (A. S. Lippolis); stefano.zoia@unito.it (S. Zoia)

5. *Limits of current computational works.* Even large language models (LLMs), the state of the art in metaphor processing, struggle to distinguish deep relational mappings from mere associations, particularly in complex or multimodal metaphors [5, 6].

Given this context, in this paper we aim to articulate a theoretical proposal that may serve both as a fruitful direction for future research and as a foundation for subsequent empirical work on this issue. Specifically, we put forward an operational framework designed to support the processing of metaphorical meaning in a way that can effectively represent and interweave both conceptual and pragmatic aspects.

The starting point for this framework lies in the observation that, for analytical purposes, implicit linguistic meaning can be treated as operating on a different level from that of explicitly encoded, surface-level meaning. Conceptual and pragmatic meanings are not directly encoded in the expression itself; rather, they are inferred through processes of contextualization and implicature. Consider, for example, the sentence *we are wasting our time*: the conceptual interpretation—where notions such as finite resources, waste, and time emerge, and where an analogical operation groups features of these concepts—is not directly expressed by the literal sentence. Similarly, the utterance's potential pragmatic function—i.e., the way that the utterance is intended—is also not explicitly encoded. This distinction reflects a view of meaning as a multi-level phenomenon—not necessarily intended as a model of how language inherently works, but as a potentially effective way of structuring meaning for computational processing—where explicitly encoded meaning and interpretation can be distinguished from conceptual and pragmatic ones, which remain implicit. Although stratified or hierarchically structured models of meaning are well established in the literature on Pragmatics (e.g. in Grice [7]), computational metaphor analysis and pragmatics have yet to establish a systematic connection—a link that could substantially advance research in both domains.

Therefore, from this theoretical perspective we propose a framework focused on the implicit level of meaning that aims to study metaphors as *cognitive tools in context*, which do not exist as an abstract operation, separate from its usage or from concrete linguistic experience. A simultaneous cognitive and pragmatic analysis would in fact support the view of metaphor as a tool that effectively works—that is, it achieves communicative efficacy—when it is perceived as conceptually appropriate and functionally aligned with the speaker's pragmatic goal, in addition to serving as models that reflect the cognitive structuring of experience, shaping both reasoning and action [1]. In this view, the implicit knowledge conveyed by metaphors can be processed more fruitfully in computational systems if conceived as a multi-layered entity, like an onion. Each of our onion layer can be ontologically represented, providing a structured and functional approach to modeling metaphorical knowledge that aligns with our analytical goals. If the outermost layer correspond to the level of contextual information mapping, moving inward reveals the level at which the analogical conceptual operations take place, and deeper still, the pragmatic intention that motivated the entire utterance. By connecting CMT's cognitive account of metaphor with SAT–inspired pragmatics, our stratified framework aim to capture both what metaphors map and what they accomplish in communication. This unified model lays the groundwork for richer datasets, standardized evaluation, and computational systems that provide a clearer account of how to unlock implicit knowledge through metaphor.

The remainder of this paper is structured as follows. In Section 2, we describe the related works. Section 3 introduces our three-dimensional model of meaning processing, illustrating how conceptual and contextual layers can be systematically integrated. In Section 4, we explore the consequences of this multi-layered view for computational processing of metaphors, discussing both implementation strategies and evaluation methods. Finally, Section 5 summarizes our contributions and outlines directions for future work.

2. Related Work

In this section, we describe the theorical background for our proposal and discuss the challenges related to existing works on metaphor processing.

2.1. Metaphor theories

CMT, developed by Lakoff and Johnson, posits that metaphors map a source domain onto a target domain via systematic correspondences, enabling abstract reasoning through familiar experiential structures [2]. Conceptual Blending Theory (CBT) extends this view by introducing a generic "blend" space that selectively inherits elements from both input domains according to a blending criterion or key property, such as *yellow* when we say "golden hair", evoking that golden is yellow and shiny like the hair [8]. CBT is regarded as a valid computational approach also by the Categorization theory [9]. This account sees metaphors as category statements where the source acquires a categorical meaning, more abstract than its literal meaning. For example, "golden" in "golden hair" denotes the category of "shiny, yellow things". In this view, conceptual blending can be used to extract the abstract meaning of the source and combine it with the meaning of the target. These frameworks emphasize that metaphor comprehension relies on shared background knowledge (or frames), which Fillmore's frame semantics formalizes by associating lexical items with structured role–filler expectations [10]. The complementary relationship between CMT, CBT, and frame semantics highlights that metaphorical meaning emerges not merely from lexical similarity but from dynamic frame activation and role alignment within a community's commonsense knowledge [4]. Beyond commonsense or prototypical knowledge, recent theories of metaphors have noted the lack of inclusion of personal and sometimes contextual aspects that influence knowledge acquisition and interchange. In fact, the experiential dimension of metaphor has traditionally been downplayed, with research focusing primarily on metaphors as a mental and individual achievement. Researchers have so far paid little attention to context and the collaborative production of metaphoric language [11].

The communicative aspect of metaphor is fundamental to view metaphor as a multidimensional phenomenon [12]. Metaphor systems are not neutral but reflect underlying belief systems that justify social actions and representations. In this view, language and metaphor in particular plays a key role in realizing these social and political values: texts are always "oriented social action" [13]. Linell's notion of an "interworld" provides a valuable theoretical framework for understanding these social dimensions of metaphor [14]. Unlike traditional cognitive approaches that locate metaphor primarily in individual minds, the interworld concept emphasizes how metaphorical meaning emerges through interaction in a shared communicative space. As an example, long-standing views of metaphor like the one carried out by CMT presuppose universal bodily experiences, excluding experiences of the disabled [15]. For this reason, recent studies claim for a view on metaphor that is not just embodied, but inter-bodily. Indeed, Gibbs [16], contrary to the standard assumption within CMT that claims source domains of conceptual metaphors are primarily based on direct sensorymotor experiences, argues that metaphorical meanings do not necessarily arise from the mappings of purely embodied knowledge onto abstract concepts. Instead, the source domains themselves metaphorical in nature.

Connected to inter-bodily multidimensional accounts of metaphor is the metaphor resistance phenomenon, only recently studied, and the various reasons why it happens. For instance, people resist metaphors if they lack explanatory power or for a preference for alternative metaphorical concepts with respect to normative ones. However, without a comprehensive metaphor study it is not possible to know why some metaphors aren't picked up [17]. Thus, we aim to shed light on these theoretical studies to account for a multidimensional view of metaphor that can also reflect in a new strand of computational metaphor processing studies.

2.2. Metaphor representation

Computational representations of metaphor have leveraged structured resources such as MetaNet and Framester, which align conceptual metaphors with FrameNet frames and roles [18]. The Amnestic Forgery Ontology further integrates MetaNet into Framester, providing a rich graph of source–target frame pairs, example sentences, and hierarchical relations among metaphors [19]. Ontological formalisms based on the Blending Ontology[1] encapsulate the four-space blending networks of CBT,

[1]Available at https://github.com/dersuchendee/BlendingOntology.

permitting explicit encoding of input spaces, generic spaces, blends, and their mapping relations. Despite these advances, existing SWRL-based and rule-driven approaches often lack scalability and fail to account for tacit, context-dependent knowledge, limiting their applicability to open-ended or multimodal metaphor interpretation [20, 21].

2.3. Metaphor datasets and benchmarks

Recent years have seen new dataset for computational metaphor processing. Among others, the VU Amsterdam Metaphor Corpus (VUA) has become a standard benchmark for metaphor detection [22]. However, such corpus only marks metaphoric tokens and does not specify the source and target domains behind each metaphor. To fill this gap, smaller domain-annotated datasets have emerged, such as the one by Gordon et al. [23], which annotates metaphorical tokens, source and target conceptual domains. A more comprehensive corpus is Metanet [24], a semantic wiki with conceptual metaphors that has been employed and extended in the Framester knowledge hub [18]. Lippolis et al. introduce the Balanced Conceptual Metaphor Testing Dataset (BCMTD), the first dataset that contains metaphors from the medical domain to test systems' generalizability [25]. At the same time, it is claimed [26] that in spite of the attention that metaphor has received over the centuries, and more recently within the cognitive paradigm, we still lack explicit and rigorous procedures for its identification and analysis, especially when one looks at authentic conversational data rather than decontextualized sentences. Furthermore, more recently, doubts have been expressed about the legitimacy of extrapolating too readily from language to cognitive structure, and distinctions have been drawn between claims about whole linguistic communities or idealised native speakers, and claims about the minds of single individuals.

2.4. Theory-driven computational processing of metaphor

Recent work in theory-driven metaphor processing integrates CMT and CBT into model architectures and annotation schemas. Mao et al. [27] and Tian et al. [28] demonstrate that embedding theoretical constraints into training objectives improves metaphor detection performance. For interpretation, unsupervised and neural methods extract source and target domains or link attributes between them [29, 30], but typically depend on single-word annotations or pre-specified targets [31]. Visual metaphor datasets such as MetaCLUE and ELCo provide multimodal challenges, yet public resources remain scarce [32, 33]. In this context, neurosymbolic systems emerge. Logic-Augmented Generation (LAG) offers a promising paradigm by treating LLMs as reactive continuous knowledge graph generators, which convert text (and images) into structured semantic graphs and then enrich them with tacit knowledge to produce extended knowledge graphs that adhere to logical constraints [34]. This hybrid approach has been explored in the work by Lippolis et al. [25], who showed that LLMs continue to struggle with metaphorical processing, especially domain specific, and multimodal inputs, despite seeing that a neurosymbolic system like LAG improves current metaphor performance.

Furthermore, Lieto et al. [35] recently presented a system called MET^{CL} able to perform metaphor generation and classification by applying a formal operationalization of the CBT. The core of MET^{CL} is a reasoning framework specialized in human-like commonsense concept combination: the Typicality-based Compositional Logic (T^{CL}) first presented in [36], which is able to account for the composition of prototypical representations. The way MET^{CL} generates metaphor representations is grounded in the Categorization theory, but the system was also applied to the conceptual metaphors from MetaNet, which is based on Conceptual Metaphor Theory.

2.5. Metaphors as speech acts

Pragmatics, as the study of the meaning of linguistic signs in context—that is, in their actual use—deals primarily with implicit linguistic knowledge: the type of meaning it investigates, the pragmatic message, is not encoded in any direct way in the literal utterance. One can say something that literally means one thing while actually intending something entirely different. For instance, saying to someone on the subway, "You're standing on my foot", means likely not wanting to describe the situation to them, but rather asking them to move. The utterance may contain only hints as to how the pragmatic meaning should be interpreted, such as a specific tone of voice, and it depends on a variety of extra-textual and extra-linguistic elements, including context, linguistic conventions, and socio-cultural norms.

Utterances that contain metaphors can, of course, be described in pragmatic terms; however, pragmatic approaches to metaphor depend on the different approaches to the interpretation of metaphor itself. In that of two leading figures in the pragmatics literature and of SAT such as Grice and Searle [37, 38] there is the idea that the metaphorical interpretation presupposes and is derived from the literal interpretation through a "decoding" of a secondary level of meaning, as if metaphors and other figures of speech were some exceptional or supplementary use of language. The assumption of a clear-cut distinction between the literal interpretation and the metaphorical interpretation of an utterance is instead abandoned in the approach to metaphor of Sperber and Wilson [39, 40], the theorists of Relevance Theory (RT). RT, whose strength lies in its understanding of how language works through its emphasis on the inferential comprehension of non-literal meanings from contextual cues and assumptions, and on the cognitive principle that human communication aims at maximal relevance with minimal processing effort, treats metaphor as a pragmatic tool conveying implicatures and implicit evaluations. The cognitive approach to the study of metaphors has been increasingly followed after the rise of CMT, and works have focused on the many and varied effects that metaphor has on cognitive processes (e.g., [41, 42]). However, in computational pragmatics the field still appears to be very open: the pragmatic processing and understanding of metaphor—why it is used, in what context, and with what effect—remains only partially addressed, despite recent advances of contextual neural models. Focusing on cognitive aspects may also obscure others, such as the performative dimension of utterances, which is emphasized in SAT. Speaking of utterances as speech acts, in fact, means considering language and linguistic utterances not merely as expressions of mental operations, like articulations of thoughts, but as actions in themselves. As actions, while pursuing their speaker's intentions, they produce effects in the world, whether intended or not, and carry out acts such as describing, ordering, pleading, or, through conventional formulas, marrying two people or sentencing someone.

A central aspect in explaining pragmatic meaning is the intention attributed to an utterance, the *illocution*—the way a speaker intends the sentence they utter, for instance as a statement or a request. One of the main goals of computational pragmatics inspired by SAT is to study how to automatically assign an *illocutionary act* to an utterance, a task framed as a problem of context dependence but complicated by several factors. First, there is the challenge of formalizing intentional, conventional, or otherwise contextual aspects that are extra-linguistic, along with all the choices that such formalization entails. Second, there is no deterministic relationship between clause types and illocutionary force: imperative clauses are not invariably commands, interrogative clauses are not always queries, and declarative clauses are not necessarily assertions. Finally, any illocutionary classification, however widely shared, will inevitably fall short of capturing the compositionality of the intentions at play in natural language. Depending on factors such as the power dynamics between speaker and addressee and their communicative goals, an utterance may emerge as a complex blend of illocutionary forces, and the same applies in the case of a "metaphorical speech act"[2]. As Michelli, Tong and Shutova [43] observe, the literature on metaphor intention remains fragmented: there is still no systematic and comprehensive account of intentions behind metaphor use, nor an operationalized framework that enables their annotation in linguistic data. Yet explaining the communicative role of metaphors in terms of taxonomies of intentions, as they suggest, risks reducing intentions to isolated categories. Moreover, although metaphor scholars do not share a common notion of intention, it is typically formalized as a prior intention—that is, a representation in the speaker's mind of their communicative goals, particularly in approaches that address metaphor through Theory of Mind models based on beliefs and intentions. As Gibbs [44] points out, conceiving of intentions as individual mental states makes them opaque, since agents are not always aware of the causes of their behaviour; this, in turn, can lead to computational abstractions that lack real explanatory value for illocutionary intention. Studying metaphors as speech acts, instead, means treating intention not as a mental state, but as a feature attributed to linguistic acts,

[2]Understood here not as a distinct illocutionary category, but as a speech act that is in some way metaphorical.

in line with the philosophy of action outlined by Austin in his linguistic analyses [45, 46]. This means that intentions are not the reasons speakers might give for using certain metaphors, but rather the intentions expressed by the utterance itself, such as whether it is interpreted as a request rather than a statement. Consequently, it is not immediately necessary to presuppose mental states beyond linguistic analysis. From a computational perspective, this entails modelling intention as an inferred property of the communication rather than as a presupposed mental condition—certainly a challenge, but also a potentially valuable contribution to computational pragmatics. If metaphor-containing utterances are speech acts, they also produce effects on the communicative context, part of which derive from metaphor as a cognitive and stylistic device. The pragmatic effectiveness of a metaphorical speech act—how well the utterance serves the communicative intention and aligns with the interlocutor's context and sensitivity—is shaped by the stylistic tone conveyed through rhetorical figures and by the strength of the conceptual evocation created by the metaphor itself. Within a given illocutionary intention, a metaphor can either reinforce or attenuate that intention: a command may sound more polite, a request more engaging, or a threat more forceful; conversely, if the metaphor is perceived as inappropriate, its expressive or persuasive force may be diminished. This suggests that, beyond illocutionary intention, it is also important to consider the *perlocutionary* level (in Austin's terms, the effects produced *by* saying something) of a metaphorical speech act, which concerns the effect an utterance or image has on the listener (e.g., convincing, frightening, provoking thought). These effects can be cognitive—often more difficult to study—but also more immediate at the emotional-psychological level, since, as figures of speech, metaphors can aim at linguistic persuasion, including the evocation of emotional responses.

3. How meaning is an onion: a proposal of three-dimensional meaning representation

3.1. The layered perspective

We now present the main thesis of our research proposal and the three-dimensional framework for processing metaphor meaning. This model provides a foundational vocabulary that, once debated within the community, can lay the groundwork for future computational implementations. Its primary aim is to offer a comprehensive interpretation of metaphor, treating it simultaneously as a cognitive instrument and as a speech act.

(a) Image of the song "Smoking kills" by Dopesmoke.[a]

[a]https://open.spotify.com/intl-it/track/6XL0aFfNf6PzhyzPddRArR

(b) Anti-tobacco campaign from the online campaign *No Smoke Revolution.*[a]

[a]https://digitaladvocacycenterwher.com/en/gun-with-cigarette-bullets/

(c) Ad by The Leith Agency for ASH UK (Dec 1994).[a]

[a]https://adsspot.me/media/prints/ash-action-on-smoking-and-health-bullet-788f341dba7b

Figure 1: Examples of anti-smoking imagery: the same metaphor can be conveyed in different contexts and with different goals.

3.1.1. Layer 1: The external context layer

The external content layer includes metadata of the communicative object under analysis: domain, provenance, connections to existing knowledge bases, etc. It also includes annotator metadata.

Consider the anti-smoking images in Figure 1. First of all, this layer would include domain classification (public health advertising), provenance references (campaign organization, publication date, media outlet where it appeared), and frame connections linking to established conceptual metaphor databases such as Metanet's HARM IS DESTRUCTION. The annotator metadata would capture the interpreter's geographical background, cultural context (attitudes toward smoking and firearms), and demographic information, recognizing that metaphor interpretation is inherently subjective and culturally situated. This foundational layer ensures that subsequent cognitive and pragmatic analyses can be properly contextualized within their social, temporal, and interpretive frameworks. Figure 2 shows an example of annotations for the metaphorical image shown in Figure 1b.

```
Domain: "Public health"
Provenance:
  • Author: "NoSmokeRevolution"
  • Media platform: "Facebook"
  • Publication date: November 15, 2013
Annotations: "A gun with cigarettes as bullets."
```

Figure 2: An example of "Layer 1" analysis for Figure 1b, showing the annotation of its content and related metadata.

3.1.2. Layer 2: The conceptual combination layer

This layer includes the cognitive context of the element, and in particular what enables the metaphorical mapping by taking into account only the two domains employed in the metaphors. Analyses that detect source and target concepts in a sentence or image are part of this layer, as well as the approaches that try to blend the two concepts to generate a representation of the metaphorical meaning. The representation of source and target can provide a rich description of the respective domains, including frame roles that describe related concepts and the semantics of their relations, including why they map (the blending principle).

The context of this layer is given at least by the annotation of source and target domains and is fully represented according to the Blending Ontology (see Section 2) with the following elements: (i) Blendable (source and target domains); (ii) Blending (the blending principle); (iii) Blended (the actual blend).

To generate a representation of the metaphorical meaning that emerges from the blending, we need a computational mechanism to realize the combination of source and target. Consider again one of the anti-smoking advertisement such as Figure 1. The source domain is SHOOTING, while the target domain is SMOKING. According to Blending Theory, the blendables are the two input spaces: the weapon frame (with roles like shooter, ammunition, target, harm) and the smoking frame (with roles like smoker, cigarettes, user, health consequences). The blending principle that enables this conceptual integration is *lethality*—both bullets and cigarettes cause death, though through different temporal mechanisms. This creates an emergent blended space where cigarettes inherit the immediate, violent danger typically associated with ammunition, while smoking adopts the intentional, direct harm associated with shooting. The blended space produces the integrated concept where each cigarette becomes a bullet the smoker fires at themselves, creating a self-destructive cycle. Frame roles map systematically: the SHOOTER maps onto the SMOKER, AMMUNITION onto CIGARETTES, the ACT OF SHOOTING onto the ACT OF SMOKING, and FATAL INJURY onto SMOKING-RELATED DISEASE, unified by the overarching principle of *lethality* that bridges the temporal gap between immediate and gradual self-harm.

Many attempts have been made to develop an algorithmic solution for conceptual combinations, each with its limitations. The Structure Mapping Engine (SME [47]) is a well-known computational approach that takes into account a broad description of source and target domains. Given the knowledge graphs representing the objects and the relations involved in the two domains, the SME finds isomorphic subgraphs that yeld the mappings. While powerful, the SME requires a rich, formal description of the two domains. A less demanding approach, directly inspired by the Categorization theory, is MET^{CL} [35]. Based on a cognitively inspired logic for conceptual combination, this system can automatically generate a prototypical representation of the metaphorical mapping (the blend) from the prototypical representations of source and target concepts. MET^{CL} consists in a three-step pipeline. The first step builds a structured representation of the metaphors to be analyzed, highlighting source and target. The second step generates a prototypical representation of both the source concept and the target concept. Each concept is represented by a prototype, i.e. a small set of typical features that can be automatically extracted from ConceptNet [48]. The third step is the conceptual combination. The T^{CL} logic combines the source and target prototypes to generate an abstract representation of the metaphor. The combinations generated by MET^{CL} were generally accepted by human judges as capturing relevant aspects of the intended metaphorical meaning.

MET^{CL} can be seen as a tool for knowledge graph completion. Indeed, manually curated resources like MetaNet are inevitably incomplete and suffer from under-representation of the wide metaphor phenomenon[3]. The ability of the system to automatically generate a representation of a given metaphor allows to cover a wider spectrum of expressions. In Lieto et al.[35], LLMs were used to classify metaphorical sentences into the MetaNet ontology classes, showing the benefits of extending the ontology with the representations generated by MET^{CL}. Our proposal for Layer 2 is to use a conceptual combination system like MET^{CL} as the reasoning mechanism. It can also be provided as a basis for the LAG-based approach by Lippolis et al. [25] (see Section 2) that addresses multimodal metaphorical raw data and generates knowledge graphs based on implicit knowledge. Figure 3 shows the MET^{CL} implementation of this layer.

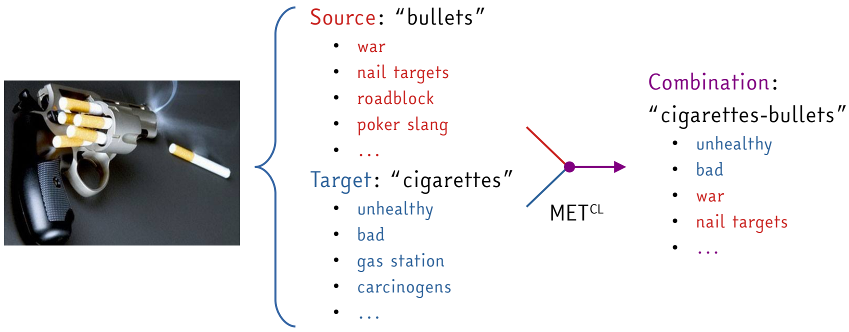

Figure 3: An example of "Layer 2" analysis for Figure 1b, showing the result of the conceptual combination performed by MET^{CL}.

3.1.3. Layer 3: The pragmatic layer

Compared with the first two layers, the third remains largely under-investigated from a computational perspective. Our aim is to discuss a categorization that would allow for its effective implementation, also with a view to the composition and annotation of a dataset. The idea, in fact, is to propose a human annotation campaign on a dataset—such as one consisting of metaphors found in images, as in examples discussed so far—and then to assess the behavior of an automatic system on the same task.

As seen in Section 2.5, works on the communicative role of metaphors in a pragmatic sense tend to explain it in terms of the intentions or discourse goals they are meant to achieve, proposing taxonomies of intentions [43], that often assumes intentions as mental states. For the analysis of intention as an illocutionary act, we reuse the standard categories of SAT, focusing on defining the role of metaphorical linguistic action[4].

[3] In particular, the MetaNet project is still under constant improvement and enrichment after more than a decade from its beginning. More information is available on the website of the MetaNet Project.

[4] The general classification of illocutionary acts (or language functions) has been a recurring topic in the philosophy of language literature, but here we are not concerned with defending a particular classification. We adopt an approach that would allow us to flexibly assess the composition of forces within an utterance containing a metaphor, focusing on the example of *directives*—the linguistic category of requests or orders typically derived from imperative sentences and found across most languages and linguistic cultures [49]. According to certain theoretical frameworks, such as those proposed by Popa-Wyatt, one could also distinguish primary and secondary illocutionary acts, sometimes nested within each other, as, for instance, in the case of an ironic act embedded inside a metaphorical one [50], but given the preliminary nature of this operational framework, we focus on annotating only the forces perceived as primary.

As discussed, a perlocutionary analysis aligns with evaluating not only what a metaphor means, but also what it does, treating utterances as actions that produce effects in the communicative context and in listeners. We propose capturing these effects through a first, immediate psychological evaluation, and extend this approach across modalities, including the recognition of tone of voice in speech as a cue to pragmatic intention.

We therefore propose the following categorization:

- **Attitude**: The speaker's evaluative stance toward the object of the metaphorical analogy (not toward individual discourse elements separately). This dimension reflects whether the attitude conveyed is positive, negative, or neutral, and can influence the tone and intention inferred from the utterance.
- **Illocutionary Act**: The act that expresses how speakers intend their utterances to be understood. Following classifications such as Searle's taxonomy, illocutionary acts can include assertive, directive, commissive, expressive, and declarative types. Communicative acts often involve a composite of multiple illocutionary forces, but our case study focuses specifically on elements that can be traced back to directive illocutionary acts as the principal pragmatic force perceived by annotators.
  - **Directive Kind**: The definition of the specific type of illocutionary act according to a range of assertive *force*. Directive acts are those in which the speaker attempts to get the addressee to perform an action, but this force can vary in intensity: they may take the form of requests, commands, suggestions, prohibitions, or pleas. We propose to model the type of directive along a continuum of assertiveness to capture this variation in illocutionary strength.
- **Perlocutionary Act**: The act concerning the effects an utterance or image has on the listener or annotator, here analyzed as psychological or emotional. This includes both intended and unintended responses, which we propose to capture as annotators' evaluations.
  - **Perlocutionary Effect**: The emotions evoked by the metaphor in the annotator, categorized according to the Emotion Frame Ontology [51]. These emotional responses may vary and are optional: in some cases a metaphor may evoke no particular emotion.
  - **Efficacy**: A measure (on a Likert scale from 1 to 5) of the metaphor's effectiveness, namely its capacity for persuasion relative to the presumed illocutionary intention, its appropriateness to the perceived context, and the annotator's personal sensitivity.
- **Other Contextual Elements**
  - **Pragmatic/Visual Clues**: Indicators that suggest the speaker's or communicator's attitude toward the metaphor's referent, e.g. use of specific colors.
  - **Pragmatic/Visual Tone of Voice**: The overall communicative tone (e.g., humorous, dramatic, ironic), which contributes to how the metaphor and its communicative intent are perceived.

As an example, consider again the image shown in Figure 1b. The speaker's evaluative stance can easily be perceived by an annotator as negative toward the analogy's object: smoking is framed as a lethal threat and self-destructive act. This negative attitude is conveyed both conceptually (via the metaphor) and visually (via design choices). The utterance, although visual, can function as a directive illocutionary act because it is embedded in the context of anti-smoking public campaigns and visually represents a cigarette as a bullet. This metaphor implies a communicative function—discouraging smoking behavior by equating it with an act of self-inflicted violence. From this function, the speaker's intention can be inferred as emerging from the pragmatic force and evaluative stance of the act itself, as to prevent smoking behavior by highlighting its deadly consequences through the conceptual mapping of SMOKING IS SHOOTING ONESELF. As for the directive kind, the discouraging attitude against smoking places the metaphorical act in a position that can be perceived by an annotator as the highly assertive end of the directive spectrum, referring to a prohibitive force similar to a command. This is achieved not through explicit verbal instruction but via the emotional and conceptual weight of the metaphorical image. The perlocutionary effects that can be experienced by an annotator can include emotions such as concern, discomfort, and shock as viewers process the visual equation of cigarettes with ammunition. The metaphor's efficacy can be perceived as potentially high due to the stark, unambiguous nature of the weapon imagery and its cultural associations with death and violence. Visual clues reinforce the negative attitude through the absence of color, and the conceptual mapping between shooting a gun and smoking a cigarette is underlined by the presence of smoke. The overall visual tone of voice is dramatic and alarming, designed to interrupt habitual thinking about smoking through visceral impact rather than rational argument. This example of analysis is outlined in Figure 4.

Attitude: *negative*
Illocutionary Act: *directive*
- Directive Kind: *highly assertive*

Perlocutionary Act: "stop smoking"
- Perlocutionary Effect: "fear", "concern", "shock"
- Efficacy: *high* (4/5)

Visual Clues: "absence of color", "smoke"
Visual tone of voice: "dramatic", "alarming"

Figure 4: An example of "Layer 3" analysis for Figure 1b, showing its description according to the pragmatic analysis.

4. Implications for computational metaphor processing and future work

Our proposal holds several implications for computational metaphor processing. In this section, we outline them and propose future research that can implement the approach into an empirical framework, going back to the five problems described in Section 1.

Datasets. Metaphor datasets should not only account for simple domain mappings within sentences, but also allow for the annotation of conversational and multimodal data, capturing dimensions beyond the purely cognitive. To date, no single metaphor-related dataset annotates attitude, speech acts and clues of utterances, although initial efforts are being made in this direction, for example in representing intentions [43]. Moreover, datasets addressing perlocutionary effects should also be linked to ontological resources on emotions. Currently, no dataset represents meaning as a spectrum, so future work should incorporate this conception of meaning. Finally, employing diverse annotators is crucial under the framework adopted in this paper, as it is necessary to establish the first gold-standard, fine-grained pragmatic datasets in the field.

Knowledge representation. Current ontological frameworks for metaphor representation, as discussed in Section 2, have made important strides by encoding source–target mappings and blend spaces. However, they tend to treat metaphor as a phenomenon abstracted from discourse and pragmatic intent. In contrast, our proposal emphasizes a multi-layered model of metaphor meaning processing that integrates contextual, conceptual, and pragmatic levels, requiring a representational shift that could, in future work, be formalized ontologically. Metaphor representation should therefore evolve toward stratified ontological models that include a pragmatic layer capturing the illocutionary and perlocutionary acts of metaphorical utterances, grounded in discourse context, communicative intent, and emotional effects. Such a layered representation would also support updates or reinterpretations of metaphorical meaning in conversational settings—for instance, through reasoning mechanisms capable of tracking metaphor reinterpretation and resistance, grounded in community norms and individual differences.

Possible application of the framework in computational metaphor processing. The proposed framework has direct implications for computational processing. By aligning processing tasks with the three dimensions, we can design systems that generate, interpret, and apply metaphors in a more context-aware, human-aligned, and socially beneficial manner. For instance, both the LAG approach [25] and the MET^{CL} [35] approach can be employed in a complementary way to assess the conceptual content and combination layer for metaphor processing. We argue that MET^{CL} and LAG are not only compatible but potentially integrable and mutually reinforcing. Specifically, LAG could leverage the prototypical property lists provided by MET^{CL} to weight RDF triples, yielding more precise definitions and reducing LLM hallucinations. Conversely, MET^{CL} could exploit LAG's capacity to identify missing frames in MetaNet or to align typical features with ontological categories, thus enabling a shared RDF layer for interoperability. For instance, LAG might detect objects in an image, while MET^{CL} could suggest which visual attributes instantiate their prototypical properties. For what concerns the pragmatic layer, the lack of datasets currently hinders computational implementation (see Section4), starting from a gold-standard annotated dataset, but LLM-based approaches such as LAG can be used to extract pragmatic features, which can in turn be analyzed.

For what concerns *metaphor generation*, current methods often rely on shallow associations or fixed templates, leading to output that lacks conceptual coherence or pragmatic fit. By incorporating the layered representation we propose, generation systems—particularly those embedded in interactive environments such as chatbots or digital companions—can produce metaphors that are not only structurally sound, but also functionally appropriate to the communicative context. Likewise, *metaphor-aware recommender systems* could translate figurative queries (e.g., "a sofa with a cozy vibe") into concrete product attributes, letting users find items that match their affective intent while keeping the system's reasoning transparent.

Metaphor understanding, too, can benefit from the layered representation we propose. By encoding both conceptual mappings and pragmatic implications, systems can disambiguate metaphorical usage more accurately, especially in real-life settings. LLMs, when coupled with structured semantic representations, may serve as effective metaphor interpreters. In such a neurosymbolic setup, the explicit surface form of expressions is parsed, mapped to conceptual domains, and then interpreted through pragmatic filters reflecting the speaker's goal, emotional tone, and discourse situation. This facilitates better performance in metaphor detection, paraphrasing, and explanation tasks, especially in underexplored genres like dialogue, narratives, or scientific texts.

Metaphors for social good. Metaphors are not neutral; they carry social, cultural and emotional weight, and using or detecting them responsibly has implications across several socially relevant domains (see Section 2), for instance employing metaphor-aware computational tools in education, science communication, doctor-patient interactions and hate speech.

5. Limitations and conclusion

Metaphors are cognitive tools deeply embedded in human reasoning and communication. In this work, we assess current challenges in metaphor processing and propose a three-layered framework that integrates both conceptual and pragmatic implicit knowledge. This onion-like stratification allows computational models to capture not only what metaphors mean but what they do, illuminating both their cognitive structure and their communicative function. In doing so, the framework aspires to bridge disciplinary boundaries, situating metaphorical analysis at the intersection of cognitive science, computational modeling, and pragmatics, thus offering a conceptual scaffolding that accommodates both ontological and epistemic variability.

Although the framework is currently theoretical, we demonstrate how it can serve as both a fruitful direction for future research and a foundation for empirical work on this issue, for instance, in designing computational systems that are metaphor-aware, context-sensitive, and socially responsible. By formalizing metaphor as a structured, multi-dimensional phenomenon, we lay the foundation for new datasets, evaluation metrics, and neurosymbolic methods that better reflect the richness of metaphor in natural discourse. This formalization highlights the interplay between conceptual structure and pragmatic function, opening avenues for interpretive flexibility in subsequent computational implementations.

The main current limitation of this framework is the lack of computational implementation. It is also necessary to develop a robust, shared line of research that integrates CBT and pragmatics theories on metaphor interpretation. This effort should include discussion of a common vocabulary for the construction of future datasets, and, on the computational side, a coherent formalization of pragmatic aspects such as intentional and contextual factors. Looking ahead, we envision extending this framework to new modalities and domains and implementing it in computational experiments.

Acknowledgments

This work was supported by the PhD scholarship “Discovery, Formalisation and Re-use of Knowledge Patterns and Graphs for the Science of Science”, funded by CNR-ISTC through the WHOW project (EU CEF programme - grant agreement no. INEA/CEF/ICT/ A2019/2063229).

Declaration on generative AI

In the preparation of this work, the authors used GPT-4 and Grammarly in order to: Grammar and spelling check. After using these tools, the authors reviewed and edited the content as needed and take full responsibility for the publication’s content.

References

[1] K. S. Moser, The role of metaphors in acquiring and transmitting knowledge, European perspectives on learning at work: the acquisition of work process knowledge (2004) 148–163.
[2] G. Lakoff, M. Johnson, Metaphors We Live By, University of Chicago Press, 1981.
[3] R. M. Hicke, R. D. Kristensen-McLachlan, Science is exploration: Computational frontiers for conceptual metaphor theory, arXiv preprint arXiv:2410.08991 (2024).
[4] B. Dancygier, Figurativeness, conceptual metaphor, and blending, in: The Routledge Handbook of metaphor and language, Routledge, 2016, pp. 46–59.
[5] T. Wijesiriwardene, R. Wickramarachchi, B. G. Gajera, S. M. Gowaikar, C. Gupta, A. Chadha, A. N. Reganti, A. Sheth, A. Das, Analogical–a novel benchmark for long text analogy evaluation in large language models, arXiv preprint arXiv:2305.05050 (2023).
[6] M. Nezhurina, L. Cipolina-Kun, M. Cherti, J. Jitsev, Alice in wonderland: Simple tasks showing complete reasoning breakdown in state-of-the-art large language models, arXiv preprint arXiv:2406.02061 (2024).
[7] H. P. Grice, Meaning, Readings in the Philosophy of Language (1971) 436–444.
[8] G. Fauconnier, M. Turner, Conceptual blending, form and meaning, Recherches en communication 19 (2003) 57–86.
[9] K. J. Holyoak, D. Stamenković, Metaphor comprehension: A critical review of theories and evidence., Psychological bulletin 144 (2018) 641.
[10] C. J. Fillmore, Frame semantics, in: Cognitive Linguistics: Basic Readings, De Gruyter Mouton, Berlin, New York, 2006, pp. 373–400.
[11] T. W. Jensen, The world between us: The social affordances of metaphor in face-to-face interaction, RASK Internationalt tidsskrift for sprog og kommunikation 47 (2018) 45–76.
[12] E. Semino, Metaphor in discourse, Cambridge University Press Cambridge, 2008.
[13] T. Titchkosky, Reading and writing disability differently: The textured life of embodiment, University of Toronto Press, 2007.
[14] P. Linell, Rethinking language, mind, and world dialogically, IAP, 2009.
[15] S. Schalk, Metaphorically speaking: Ableist metaphors in feminist writing, Disability Studies Quarterly 33 (2013).
[16] R. W. Gibbs, Metaphor wars, Cambridge University Press, 2017.
[17] R. W. Gibbs Jr, J. Siman, How we resist metaphors, Language and Cognition 13 (2021) 670–692.
[18] A. Gangemi, M. Alam, L. Asprino, V. Presutti, D. R. Recupero, Framester: A wide coverage linguistic linked data hub, in: Knowledge Engineering and Knowledge Management: 20th International Conference, EKAW 2016, Bologna, Italy, November 19-23, 2016, Proceedings 20, Springer, 2016, pp. 239–254.
[19] A. Gangemi, M. Alam, V. Presutti, Amnestic forgery: An ontology of conceptual metaphors, in: S. Borgo, P. Hitzler, O. Kutz (Eds.), Formal Ontology in Information Systems - Proceedings of the 10th International Conference, FOIS 2018, Cape Town, South Africa, 19-21 September 2018, volume 306 of *Frontiers in Artificial Intelligence and Applications*, IOS Press, 2018, pp. 159–172. URL: https://doi.org/10.3233/978-1-61499-910-2-159. doi:10.3233/978-1-61499-910-2-159.
[20] K. Hamilton, Towards an ontology for propaganda detection in news articles, in: The Semantic Web: ESWC 2021 Satellite Events, 2021, pp. 230–241. URL: https://doi.org/10.1007/978-3-030-80418-3_35. doi:10.1007/978-3-030-80418-3_35.
[21] J. Mitrović, C. O’Reilly, M. Mladenović, S. Handschuh, Ontological representations of rhetorical figures for argument mining, Argument & Computation 8 (2017) 267–287. URL: https://doi.org/10.3233/aac-170027. doi:10.3233/aac-170027.
[22] T. Krennmayr, G. Steen, Vu amsterdam metaphor corpus, Handbook of linguistic annotation (2017) 1053–1071.
[23] J. Gordon, J. R. Hobbs, J. May, M. Mohler, F. Morbini, B. Rink, M. Tomlinson, S. Wertheim, A corpus of rich metaphor annotation, in: Proceedings of the Third Workshop on Metaphor in NLP, 2015, pp. 56–66.

[24] E. Dodge, J. Hong, E. Stickles, O. David, The metanet wiki: A collaborative online resource for metaphor and image schema analysis, in: 12th International Cognitive Linguistics Conference (ICLC 12), 2013.
[25] A. S. Lippolis, A. G. Nuzzolese, A. Gangemi, Enhancing multimodal analogical reasoning with logic augmented generation, arXiv preprint arXiv:2504.11190 (2025).
[26] E. Semino, J. Heywood, M. Short, Methodological problems in the analysis of metaphors in a corpus of conversations about cancer, Journal of pragmatics 36 (2004) 1271–1294.
[27] R. Mao, X. Li, H. Kai, M. Ge, E. Cambria, Metapro online:: A computational metaphor processing online system, in: Proceedings of the 61st annual meeting of the association for computational linguistics, Association for Computational Linguistics (ACL), 2023.
[28] Y. Tian, N. Xu, W. Mao, A theory guided scaffolding instruction framework for llm-enabled metaphor reasoning, in: Proceedings of the 2024 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (Volume 1: Long Papers), 2024, pp. 7731–7748.
[29] E. Shutova, Annotation of linguistic and conceptual metaphor, Handbook of linguistic annotation (2017) 1073–1100.
[30] Z. Rosen, Computationally constructed concepts: A machine learning approach to metaphor interpretation using usage-based construction grammatical cues, in: Proceedings of the workshop on figurative language processing, 2018, pp. 102–109.
[31] L. Wachowiak, D. Gromann, Does gpt-3 grasp metaphors? identifying metaphor mappings with generative language models, in: Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), 2023, pp. 1018–1032.
[32] A. R. Akula, B. Driscoll, P. Narayana, S. Changpinyo, Z. Jia, S. Damle, G. Pruthi, S. Basu, L. Guibas, W. T. Freeman, et al., Metaclue: Towards comprehensive visual metaphors research, in: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, 2023, pp. 23201–23211.
[33] Z. Y. Yang, Z. Zhang, Y. Miao, The elco dataset: Bridging emoji and lexical composition, in: Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024), 2024, pp. 15899–15909.
[34] A. Gangemi, A. G. Nuzzolese, Logic augmented generation, Journal of Web Semantics 85 (2025) 100859. URL: https://www.sciencedirect.com/science/article/pii/S1570826824000453. doi:https://doi.org/10.1016/j.websem.2024.100859.
[35] A. Lieto, G. L. Pozzato, S. Zoia, The Delta of Thought: Channeling Rivers of Commonsense Knowledge in the Sea of Metaphorical Interpretations, in: Proceedings of IJCAI 2025, 34th International Joint Conference on Artificial Intelligence, AAAI Press, Montréal, Canada, 2025. URL: https://github.com/StefanoZoia/METCL/blob/main/IJCAI__25_Lieto_Pozzato_Zoia.pdf.
[36] A. Lieto, F. Perrone, G. L. Pozzato, E. Chiodino, Beyond subgoaling: A dynamic knowledge generation framework for creative problem solving in cognitive architectures, Cognitive Systems Research 58 (2019) 305–316. URL: https://linkinghub.elsevier.com/retrieve/pii/S1389041719304632. doi:10.1016/j.cogsys.2019.08.005.
[37] H. P. Grice, Logic and conversation, in: D. Davidson (Ed.), The logic of grammar, Dickenson Pub. Co., 1975, pp. 64–75.
[38] J. R. Searle, Metaphor, in: A. Ortony (Ed.), Metaphor and Thought, Cambridge University Press, 1993, pp. 83–111.
[39] D. Sperber, D. Wilson, Relevance: Communication and Cognition, Blackwell, Oxford, 1986/1995.
[40] D. Sperber, D. Wilson, Representation and relevance, Mental (1988).
[41] G. Steen, The paradox of metaphor: Why we need a three-dimensional model of metaphor, Metaphor and Symbol 23 (2008) 213–241. doi:10.1080/10926480802426753.
[42] G. J. Steen, Thinking by metaphor, fast and slow: Deliberate metaphor theory offers a new model for metaphor and its comprehension, Frontiers in Psychology Volume 14 - 2023 (2023). URL: https://www.frontiersin.org/journals/psychology/articles/10.3389/fpsyg.2023.1242888. doi:10.3389/fpsyg.2023.1242888.
[43] G. Michelli, X. Tong, E. Shutova, A framework for annotating and modelling intentions behind metaphor use, arXiv preprint arXiv:2407.03952 (2024).
[44] R. Gibbs, Intentions in the Experience of Meaning, Cambridge University Press, 1999. URL: https://books.google.it/books?id=FD1rVU4LtBgC.
[45] J. Austin, How to Do Things with Words, A galaxy book, Harvard University Press, 1962.
[46] J. L. Austin, A plea for excuses, in: J. L. Austin (Ed.), Philosophical Papers, 1961.
[47] B. Falkenhainer, K. D. Forbus, D. Gentner, The structure-mapping engine: Algorithm and examples, Artificial Intelligence 41 (1989) 1–63. doi:10.1016/0004-3702(89)90077-5.
[48] R. Speer, J. Chin, C. Havasi, ConceptNet 5.5: An Open Multilingual Graph of General Knowledge, Proceedings of the AAAI Conference on Artificial Intelligence 31 (2017). URL: https://ojs.aaai.org/index.php/AAAI/article/view/11164. doi:10.1609/aaai.v31i1.11164.
[49] S. C. Levinson, Pragmatics, Cambridge University Press, 1983.
[50] M. Popa-Wyatt, Compound figures: Priority and speech-act structure, Philosophical Studies 174 (2017) 141–161. doi:10.1007/s11098-016-0629-z.
[51] S. De Giorgis, A. Gangemi, Efo: the emotion frame ontology, arXiv preprint arXiv:2401.10751 (2024).