**Metaphor Identiﬁcation Procedure**:

Abridged version of the MIPVU for human annotators[[1]](#footnote-1)

As a general rule, the annotation procedure follows the steps below:

1. Read the entire text/ discourse to establish a general understanding of the meaning.
2. Determine the lexical units in the text/discourse.
3. a. For each lexical unit in the text, establish its meaning in context, i.e. how it applies to an entity, relation or attribute in the situation evoked by the text (contextual meaning). Take into account what comes before and after the lexical unit.

b. For each lexical unit, determine if it has a more basic contemporary meaning in other contexts than the one in the given context. For our purposes, basic meanings tend to be:

* more concrete/ physical: e.g. what they evoke is easier to imagine, see, hear, feel, smell, and taste;
* related to bodily action;
* human-oriented (e.g. in terms of metaphors based on personification)
* more precise (as opposed to vague);
* historically older.

However, historical metaphor is not identiﬁed as metaphorical by MIP. For instance, the words *fervent* and *ardent* used to have two senses, one for temperatures and one for emotions. However, in contemporary British English both terms have lost their original temperature sense: in the Macmillan dictionary\*, for instance, they only have their present-day emotion senses. Hence expressions like *ardent lover* are not judged to be metaphorical when analyzed by MIP because there is no contrast between the contextually appropriate emotion sense and the historically older and more basic temperature sense: the latter is simply not available to the typical contemporary language user anymore, as is reﬂected by the descriptions of the words in the modern users’ dictionary.

It is also worth mentioning that basic meanings are not necessarily the most frequent meanings of the lexical unit.

c. If the lexical unit has a more basic current/contemporary meaning in other contexts than the given context, decide whether the contextual meaning contrasts with the basic meaning but can be understood in comparison with it.

4. If yes, mark the lexical unit as metaphorical (with a # hashtag, in our case).

Contrary to common practice in cognitive linguistics, the MIP procedure does not aim to identify the precise nature of the underlying conceptual mappings between domains, such as *argument* and *war*, or *emotions* and *temperatures*, themselves. They identify the linguistic forms of metaphor, not its conceptual structures. In order to identify a word or set of words as metaphorically used, it is often sufﬁcient to be able to say that there are two senses and that they may be related by comparison, or nonliteral similarity. For consistency, the meanings of lexical units should be looked up in the Macmillan Dictionary (<https://www.macmillandictionary.com/>)\*.

In this study, we use the level of the word, or **lexical unit**. As a result, in order to consistently measure metaphor at one level of usage, lexical units need to be systematically and exhaustively examined for metaphorical use, and annotated as such. Here, lexical units are basically word classes, not lemmas. We analyze by word class because word classes have the closest connections with conceptual and referential classes like entities, processes, and attributes.

More specifically,

a. the contextual meaning of nouns, verbs, adjectives, adverbs, prepositions, and interjections cannot be compared with the meaning of other word classes for the same lemma (conversions); for instance, the meaning of *shift* as a noun should be analyzed irrespective of the meaning of *shift* as a verb;

b. the contextual meaning of verbs used as linking verbs, primary verbs, modal verbs, verbs initiating complex verb constructions such as start, stop, continue, quit, keep, and so on, causative verbs (have, get, and so on), and full verbs cannot be compared with the meaning of the same verbs used in other roles;

c. the contextual meaning of verbs used transitively can as a rule not be compared with the meaning of the same verbs used intransitively;

d. the contextual meaning of nouns used to designate countable entities can as a rule not be compared with the meaning of the same nouns used to designate uncountable entities.

All so-called **polywords** in the corpus are taken as single lexical units. Examples include *a good deal*, *by means of*, and *of course, a bit, a lot, at least, at all, as well, sort of/ kind of, as long as, in part*. We should **not** examine the parts of these polywords for potential metaphorical meaning.

Rather counterintuitively though, some phrases are considered and coded as consisting of separate lexical units because they have alternatives for non-metaphorical usage, like “in my medical practice” versus “#in practice, it is rather complicated”, or have no synonymous one word equivalents (compare to*of course/certainly*). Some instances of such cases include #by the #way, #on average, #to some #extent, #in #this respect, #in a #way, #in other words, #on Sunday, etc.

When annotating **proverbs**, we usually mark all the content words in the expression. For example, “not #enough #room to #swing a #cat”.

Apart from direct and indirect metaphors coded for, there is another class that expresses cross-domain mappings implicitly. Here is an example from the BNC: ‘Naturally, to embark on such a *step* is not necessarily to succeed immediately in realising *it*’. Here *step* is related to metaphor, and *it* receives a code for (implicit) metaphor. This idea may need some explication. In discourse analysis, the discourse would have to show the previous concept (antecedent, *‘step’*) instead of the cohesive element, *it*, and this would make the current proposition containing the cohesive element metaphorical. But the language in the surface text would be implicitly metaphorical, because the language does not signal the need for nonliteral comparison, as is the case with indirect and direct metaphor. Instead, implicit metaphor is due to the underlying cohesive link (grammatical and/or semantic) in the discourse which points to recoverable metaphorical material. Thus, we would code ***it*** here as a metaphor. **Pronouns** referring to the metaphorical lexical unit also gets coded as a metaphor. In principle, it is possible for both **demonstratives** as well as general words such as *thing* and *stuff* to refer back to a metaphorically used expression. In that case, they are all annotated.

In the same vein, speech participants often refer anaphorically to previously shared topics within the conversation with *that* and *this* replacing the full referent. This creates a sense of cohesion within the text, e.g. “Oh no I just wanted to know if you wanted any *that* was all”. In this example, the word *that* in “that was all” refers back to the antecedent ‘I just wanted to know if you wanted any’. Since basic and contextual meaning of “that” is contrasted here, we are dealing with a metaphorically used word. Conversely, *that* in expressions like “that’s it” and “that’s right” is not marked when not used anaphorically.

**Phrasal verbs** are verbal expressions consisting of more than one word, such as *look up* or *turn on*. We should treat all phrasal verbs as single lexical units: their individual parts do not require independent analysis for potential metaphorical meaning. The phrasal verb as a whole, however, can still be used metaphorically. For instance, *setting up* an organization is a metaphorical variant of *setting up* a road block. The problem with phrasal verbs is their superﬁcial resemblance to prepositional verbs (e.g. *based on*) and to verbs followed by free adverbs. The latter two cases should be analyzed as free combinations consisting of two independent lexical units, as opposed to phrasal verbs which should be taken as only one.

Some cases might be confusing. An example is *look up* in a sentence like “she looked up into the sky”. The contextual meaning — “to direct your eyes towards someone or something so that you can see them” — is not one of the meanings of the phrasal verb (unlike, for instance “to try to ﬁnd a particular piece of information”). The contextual meaning, instead, is the result of a free combination of a verb plus an adverb; the words consequently have to be analyzed as two separate lexical units in this sentence.

Additionally, there is the matter of complex phrasal verbs, such as *make up for* or *do away with*. These may be easily confused with combinations of simple phrasal verbs with a preposition (*make up + for* or *do away + with*). However, we are to take the dictionary classiﬁcation of these complex verbs as single units as our guideline.

In other words, in our annotations we would code “*to #make #up #for”* (all constituents) as one lexical unit (a complex phrasal word) but we would code “*to differ #on”* as two separate units (a prepositional verb), one of which may be used metaphorically but other parts may not.

**Compounds** are single lexical units consisting of two distinct parts, which may cause orthographical problems. They can be spelt in three ways: as one word, as two hyphenated words, and as two separate words. When a compound noun is spelt as one word, such as underpass, and can be found as such in the dictionary we treat it as one lexical unit designating one referent in the discourse. When a compound noun is spelt as two hyphenated words and can be found as such in the dictionary, such as pitter—patter, we similarly treat it as one lexical unit. However, if we are dealing with a novel formation unknown to the dictionary, the compound noun is analyzed as two separate units. This also applies to hyphenated compound nouns created through a productive morphological rule but that are not listed as a conventionalized compound in the dictionary (such as “under-ﬁve”).

In **spoken discourse**, nicknames, terms of endearment and swearwords, for example, are often metaphorical, as are units like “thing”, “stuff” etc. We can also see a lot of examples of metaphoric phrasal words like “#come #on, you’ve got to be kidding”, or fixed expressions like “#there we are”.

We should also identify all instances where the basic meaning refers to space, while contextual meaning refers to time. For example, *past* in “twenty #past seven” would be coded, as well as *before* in “#before I say something”.

Finally, the general rule used in annotations is ‘**When In Doubt, Leave It In**’ (WIDLII). Therefore, in dubious cases, we would annotate lexical units that are potentially metaphorical. For example, in the expression “to keep an eye on”, we would mark all the signiﬁcant words (#keep, #eye, #on) that may or may not be metaphorically used (e.g. an eye might be actually needed in tracking what others are doing) as potentially metaphorically used.

1. <https://books.google.ch/books/about/A_Method_for_Linguistic_Metaphor_Identif.html?id=lrc0-OXtnA0C&redir_esc=y> [↑](#footnote-ref-1)