

Metaphor-Guidelines

27, August 2015

Pechlivanis Konstantinos (*email*: kpechlivanis21@gmail.com)

Eirini Florou

What is a domain?

The universe of documents is characterized by a wide variety of typologies, which correspond to different layouts and logical structures. A domain can be defined as a group of documents which can be clustered with respect to the subject, for instance journals, papers, business letters are different domains. Each domain can also be characterized by some features which can make effective the domain classification. Such features can be concerned with the vocabulary of the text, its terminology and the common themes, which may appear at the texts of the same domain. To sum up, as domain we can consider a wider subject field which includes a variety of texts with similar topics, terminology and vocabulary.

What is not a domain?

It's commonly accepted that there is a part of texts that may appear some common elements, although, they don't belong to the same domain. This occurs as texts with different layout belong to different domains. Moreover, texts with different terminology can't be included to the same domain. Furthermore, a domain doesn't exist if there aren't terms from a common, specific scientific field.

Document domain

The amount of available electronic data is increasing rapidly. So, for the purposes of the Text Mining, it is necessary to manage that data. The text classification is sometimes a hard task but it is very useful in order to extract useful information about texts with relevant features. For this purpose, it is necessary to be primarily determined the subject or the topic of the document. This happens as writing on a certain subject, a certain set of typical words tend to be used. In other words, texts of similar topics will use a similar set of topical words, while texts without similar topics will use far fewer of these words.

Apart from the vocabulary, some specific terms are more likely to appear for a certain topic than for other topics. Simultaneously, other terms are far more generic and will appear in almost every topic regardless of how similar the topics may be. We should have in mind that the same term may be present in both domains.

Detecting the Domain of a Document according to IPTC

The International Press Telecommunications Council (IPTC)¹ creates and maintains sets of concepts in order to be assigned as metadata values to news objects like texts. This allows for a consistent coding of news metadata across news providers and over the course of time. The set of concepts is called taxonomy and it is a set of terms to express a facet of news content. Facets could be for example the subject, the genre or even the urgency. A taxonomy could be a flat list of terms or a hierarchical structure.

The news metadata are grouped into categories basing on content, Media Topics, genres and world regions. Especially, the genre indicates a nature, journalistic or intellectual characteristic of an item.

According to the IPTC, there are 17 special topics², that can be classified the Media articles. Especially, these topics are:

- Arts, culture & entertainment
- Crime, law & justice
- Disaster & accident
- Economy, business & finance
- Education
- Environment
- Health
- Human interest
- Labour
- Life style & leisure
- Politics
- Religion & belief
- Science & technology
- Society
- Sport
- Conflicts, war & peace
- Weather

Bearing in mind the previous distinction, after gathering the articles which will be the train corpus, we tried to classify them to taxonomies according to the prototype of the IPTC.

¹www.iptc.org/site/Home/

²<http://show.newscodes.org/index.html?newscodes=medtop&lang=en-GB&startTo=Show>

What is a metaphor?

A metaphor is a figure of speech in which an implied comparison is made between two unlike things that actually have something important in common. We can say that two parts of speech are comparing or contrast each other. This allows us to use fewer words and forces the reader or listener to find the similarities. The word metaphor itself is a metaphor, coming from a Greek word meaning to “transfer” or “carry across”. Metaphors “carry” meaning from one word or idea to another.

The simplest form of metaphor is: “The [first thing] is a [second thing].” Let’s see the following example:

- Her **home** was a **prison**.

In the above sentence, we understand immediately that her home had some of negative characteristics. Mainly, we imagine, she could not leave her home. She was trapped inside. In this sentence, “prison” is a metaphor. Look at another example:

- **George** is a **sheep**.

What is one characteristic of sheep? They follow each other. So we can imagine that George is a follower, not a leader. In this sentence “sheep” is a metaphor.

The above metaphors have the simplest forms as they are nouns. But there are other ways of making metaphors, for example with verbs or adjectives.

Metaphor example	Original sense of the word (example)
The committee shot her ideas down one by one.	Anti-aircraft guns shoot down planes.
He broke into her conversation.	Burglars break into buildings.
Tim lost his job after a heated argument with his boss.	We have a heated swimming pool.
He was dressed rather vulgarly in a loud checked suit.	I can’t hear you because the radio is too loud.

Especially, in the case of Greek language there is a variety of words which may appear either in a literal or in a non literal context.

Literal example	Non literal example
Γλυκιά καραμέλα	Γλυκιά ζωή
Το βάρος του μπαούλου	Το βάρος των ετών
Η γυναίκα γέννησε το μωρό	Το μυαλό του γεννά ιδέες
Κάρφωσε το καρφί	Κάρφωσε τα μάτια του
Σηκώνω βάρος	Δε σηκώνω αστεία
Τρώω φαγητό	Έφαγα τον κόσμο να σε βρω
Κολυμπούσε στη θάλασσα	Κολυμπούσε στον ιδρώτα

What is not a metaphor?

A metaphor is sometimes confused with a simile which compares two subjects using “like” or “as”. An example of simile would be: “He was as sly as a fox”. While a metaphor would be “He was a fox”. Within this work, we will not take into account either simile or metonymy, as we are going to concentrate entirely on metaphor detection. Metonymy enables us to use one part or aspect of an experience to stand for some other part (or the whole) of that experience. Some typical types of metonymy are:

- Part for Whole e.g. Many hands make light work.
- Whole for Part e.g. Australia beat Canada at cricket.
- Place for Institution e.g. The White House isn’t saying anything.
- Producer for Products e.g. I like Shakespeare most.
- Potentiality for Actuality: A potential event (e.g. the ability, possibility, permission, obligation to undertake an action) is metonymically linked to its actual occurrence, for example, She was able to finish her dissertation.

Unlike metaphor which involves two domains of experience, metonymy only requires one. Unlike metaphor which is based on similarity, metonymy requires contiguity.

Apart from simile or metonymy we should distinct the metaphor and the delexical or light verbs. “Delexical verb is a verb which has very little meaning in itself and is used with an object that carries the main meaning of the structure”³. Moreover, the delexical verbs have a corresponding single verb (e.g. have a listen – to listen). Especially, in Greek language there is a variety of delexical verbs like *χάνω*, *βάζω*, *παίρνω*, *δίνω*. These verbs are used with an object replacing a single verb in order to declare the action or the state. For example, may appear the delexical verbs *χάνω ερώτηση* or *δίνω υπόσχεση*, although there are the single verbs *ρωτώ* and *υπόσχομαι*, correspondingly. We should refer that, although we discriminate the delexical verbs with the metaphor expressions, we want to be annotated the delexical verbs as a separated category.

If there is a name entity, like *Χρυσή Αμμουδιά*, which is simultaneously a metaphor expression, we don’t annotate it as metaphor, although it is, as this metaphor is the name of a village.

In the phrase “to grasp the concept” the physical action “to grasp” is used as a metaphor for “to understand” (which is non-physical). But this phrase has been used so often that most English speakers do not have an image of the physical action in their mind. This metaphor has died; it is a “dead metaphor”. Respectively, in the case of Greek language, there are cases like *«υποστηρίζω την άποψη, πόδι τραπεζιού»* where neither the speaker nor the hearer can understand the non literal meaning of the phrase.

Finally, it could be another one type of metaphors “the implicit metaphor”. Implicit metaphor is due to an underlying cohesive grammatical and/or semantic

³Delexical verbs are treated in this works as defined by Collins Cobuild Grammar (1990)

link in the discourse which points to recoverable metaphorical material. For example “Naturally, to embark on such as step is not necessarily to succeed in realizing it.” In principle, it does not call for a non-literal indirect or direct comparison. Note, however, that it refers back to the metaphorically used lexical unit step. Since an analysis of the discourse would need to show step instead of it, it becomes implicitly metaphorical. Although, this type of metaphors is very frequent, does not include in the material that is investigated in the merge of the present work.

Kinds of metaphor

Humans often use metaphor to describe abstract concepts through reference to more concrete or physical experiences. Metaphorical expressions may take a great variety of forms, ranging from conventional metaphors, which we produce and comprehend every day to poetic and novel ones. Rhetorical theorists and other scholars of language have discussed numerous dimensions of metaphors. However, we are going to concentrate on detecting four kinds of metaphor:

1. indirect, lexical metaphor: metaphor at the level of a single word sense (παροιμιώδης οργάνωση, παροπλισμένοι δημοσιογράφοι, μαγικά νησιά, ομηρικοί καυγάδες). A term identifies metaphorically another literal term. The literal term may be a single term or even a subordinate nominal clause (υπογράμμισε πως η κυβέρνηση θα έπρεπε να έχει ζητήσει εξηγήσεις).
2. multi-word metaphorical expression (ανοίγω την καρδιά μου, έβαλε το χέρι του, κόβει ταχύτητα, άφησε το στίγμα του, έβαλε φρένο). A new meaning becomes due to all the constituents of the phrase. However, each constituent may be replaced by another term with similar meaning.
3. idiomatic metaphorical expressions (e.g. kick the bucket, rock the boat, put a damper on, χρούω τον κώδωνα του κινδύνου, λυδία λίθος, φωνή βοώντων εν τη ερήμω, στο μάτι του κυκλώνα, επί ξύλου κρεμάμενος). The certain expression consists entirely of the specific constituents and no one of them can be replaced by another word with similar meaning. In the Appendix A, there are more typical metaphorical expressions.
4. direct, IS-A metaphor (e.g. Her home is a prison, ο χρόνος είναι χρήμα). The core of these expressions is a copula that connects the subject of the certain verb with its complement. In the Appendix B, there are typical cases of copula.

The metaphor and the document in which appears

According to our research proposal the document’s domain is different from the domain of metaphors which appear at the certain document. In order to investigate this case, we are going to construct a corpus with texts from a range of domains. The texts have been selected from a bigger corpus which includes articles from newspapers. Then, we are going to annotate the metaphors of all the texts. For each metaphor we will annotate its type and we will select its domain. Specifically, we have already found 10 thematical domains. The metaphors can be classified into them. For instance, the metaphorical expression «οι εκλογές συνθέτουν το σκηνικό» is typical example of the second type of metaphors. The metaphor’s domain is “Arts, culture & Entertainment” while the document’s domain is “Politics”.

In order to annotate the type of metaphor we use the labels “type 1”, “type 2”, “type 3” and “type 4”, which correspond to the four kinds of metaphor that have described above. After the type annotation, we have to select the metaphor’s domain. In the case of the “type 3”, as this kind’s metaphors are idiomatic expressions isn’t necessary to annotate the domain of the metaphor and the text since the certain expression may appear with the same structure at every kind of text.

We should have in mind that may appear expressions which have a non literal meaning and they will belong to the same domain with the whole text. Moreover, may appear an expression with a non literal meaning at the certain context while the same expression at another context will have a typically literal meaning.

Document Domains

The annotator has the choice to select between 10 types of document domain. The distinction of document domains was done according to IPTC. It is necessary to have read the whole text in order to decide about its type. The annotator should have in mind that it's possible a document to discuss not only a topic but two or more subjects. In this case, it is required to be selected the part of the text which deals with each topic and to be annotated separately the topic of each part of the whole text. Especially, the domain of a document (or the domain of each part of the whole document) may be:

- **Arts, culture & entertainment:** Matters pertaining to the advancement and refinement of the human mind, of interests, skills, tastes and emotions. Especially, the subcategory of arts and entertainment deals with the collective expression of message or emotion through fine or performing arts, possibly for the amusement or relaxation of people. This subcategory could be included texts related to animation, cinema, film festival, dance, fashion, literature, fiction, poetry, music, theatre, painting, photography, sculpture, art exhibition. The subcategory of culture could be included texts associated with the ideas, customs, arts and skills of a particular group. So, customs and tradition, festive event, language, monument could be the topics of the articles which are classified at the certain category. We should refer that this category includes texts which deal with mass media and newspaper.
- **Crime, law & justice:** Establishment and /or statement of the rules of behaviour in society, the enforcement of these rules, breaches of the rules and the punishment of offenders. Organisations and bodies involved in these activities. This domain includes five sub- domains:
 1. Crime: arson, assault, computer crime, corporate crime, corruption, drug related crimes, fraud, hijacking, homicide, kidnapping, organised crime, terrorism, theft, war crime
 2. Judiciary: court, out of court procedures like arbitration, and prosecution
 3. Justice and rights: civil rights
 4. Law: civil, criminal and international law
 5. Law enforcement: arrest, criminal investigation, police, criminal punishment
- **Disaster and Accident:** Man made and natural events resulting in loss of life or injury to living creatures and/or damage to inanimate objects or property. Any unplanned event that causes unwanted consequences. An accident involving one or more vehicles. Accidents involving craft in air or space. Accidents involving marine vessels or trains. Accidents on roads. Serious or sudden misfortune like famine, fire, drought, earthquake, wind storms, volcanic eruption. A sudden, unexpected event that requires immediate action to prevent serious consequences like structural failures, transport incident, air or space or maritime or railway or road incident. Planning for actions to deal with sudden, unplanned events. Organisations set up to provide help to those in need because of lack of food, water or shelter.

- **Economy, business & finance:** All matters concerning the planning, production and exchange of wealth. The certain domain includes four sub-domains:
 1. Business information: business finance, human resources, strategy and marketing
 2. Economic sector: agriculture, chemicals, computing and information technology, construction and property, consumer goods, energy and resource, financial and business service, manufacturing and engineering, media, metal and mineral, process industry, tourism and leisure, transport
 3. Economy: This sub-domain is separated into two factors:
 - (a) economic policy: nationalisation, privatisation and state-owned enterprise
 - (b) macro economics: bonds, budgets and budgeting, business enterprise, central bank, consumers, credit and debt, currency value, deflation, economic growth, economic indicator, economic organisation, emerging market, employment statistics, exports, government aid, government debt, imports, inflation, interest rate, international economic institution, international trade, investments, money and monetary policy, mortgage, mutual funds, prices, recession, tariff
 4. Market and Exchange: which can be divided into six other factors:
 - (a) commodity market: energy market, metal, soft commodity
 - (b) debt market
 - (c) foreign exchange market
 - (d) loan market
 - (e) securities
 - (f) stocks
- **Education:** All aspects of furthering knowledge of human individuals from birth to death. The domain of the Education has five sub-domains:
 1. parent organisation
 2. religious education
 3. school: elementary schools, further education, high schools, higher education, middle schools, preschool
 4. social learning
 5. teaching and learning: curriculum, examination, students, teachers
- **Environment:** All aspects of protection, damage, and condition of the ecosystem of the planet earth and its surroundings. Climate change. Global warming including temperature research, remote sensing on temperature trends, debate on global warming, ways to reduce emissions and carbon trading. Preservation of wilderness areas, flora and fauna, including species extinction. Conservation of electrical, and other power sources. Areas set aside for preservation.

Environmental pollution like corruption of air, water, land etc. by harmful substances. Environmental issues related to the exploitation of natural resources. Open areas of trees either available for public enjoyment, or for commercial purposes. Elevated land masses formed over the ages either by erosion, volcanic eruption, or movement of massive geographical formations called plates. People and their growth and development within a natural setting. The environmental impact of renewable energy, including solar, wind, hydro, biomass and geothermal. Environmental issues about bodies of water, including oceans, lakes, streams and reservoirs, as well as ice, glaciers and forms of precipitation. Salt water masses separating continents or other major geographical masses. Smaller forms are seas or lakes or ponds. Moving water areas bounded by land that extend from earth sources and meander through land areas to join with other water areas. In smaller forms they are creeks, rivulets, streams etc. Areas generally marshy and not either under water or dry land. Often related to aquifers for water quality and/or wildlife. The natural world in its entirety. A system of plants, animals and bacteria inter-related in its physical/chemical environment. The endangered species and the invasive species.

- **Health:** All aspects pertaining to the physical and mental welfare of human beings, treatment of diseases and other conditions to produce wellness, physical fitness. Any harmful or destructive health condition affecting humans. Non-specific maladies such as stomach aches, headaches, flat feet, lumbago. A traumatic condition of the human body caused usually by outside forces. Illnesses and disorders affecting the mind and behaviour. Facilities used for any kind of healthcare. Health organizations and health treatment. Diet and dietary supplements. Medical drugs and medicine. Physical fitness and preventative medicine. Healthcare policy and medical profession. Non human diseases like any harmful or destructive health condition affecting animals, plants or micro-organisms.
- **Human interest:** Items about individuals, groups, animals, plants or other objects with a focus on emotional facets. Achievements by individuals or groups, animals, plants or other objects, such as winning a competitive contest etc. Decorations, awards and prizes, including those given by the government. Non-sport achievement by an individual or group that sets a new record, such as most books read, holding your breath longest. Rituals of a non-religious nature such as dedications, wreath laying, ship launchings etc. Individuals or groups with a focus on emotional facets. Life and behaviour of well-known people. Life and behaviour of famous people. Imperial and royal family members in their societal rather than governmental roles.
- **Labour:** Social aspects, organisations, rules and conditions affecting the employment of human effort for the generation of wealth or provision of services and the economic support of the unemployed. Working for an organisation, an individual or being self-employed, including voluntary work. People who are in the work force at a learning level, serving under a master. Child labour: Employment of children, usually under the age of 18. A person who trades

his time and talent for money and other benefits. The person who employs workers and pays the salaries and other benefits to workers. Learning skills required for employment. Providing a worker training in new skills when his old skills are no longer applicable. Generally one's trade, profession or business although professions are usually such things as lawyers and doctors, and occupations are such as clerks or truckers. Employee compensation. Benefits provided to an employee by a company, such as pension contributions, health insurance. Money paid by individual and companies into a government account to be paid out when that individual reaches a certain age. Laws governing employment. Rules and procedures for guaranteeing the health of workers. Strike. Retirement. Pension. Unemployment. Job layoffs. Groups established to represent bodies of workers.

- **Lifestyle and leisure:** Activities undertaken for pleasure, relaxation or recreation outside paid employment, including eating and travel. Organisations joined by individuals because of similar interests. Contests generally for one's amusement. Board game, card game, children's game, dice game, outdoor game, puzzle, tile game, video game, gaming and lottery, hobby, holiday, leisure venue, amusement park, bar, cafe, nightclub, restaurant, sport facilities, recreational activities, fishing, hunting, mountaineering, travel, tourism. Interest in houses, apartments and their decoration, including furnishings. Interest in developments in style, fashion and design. Social event with dancing.
- **Politics:** Local, regional, national and international exercise of power, or struggle for power, and the relationships between governing bodies and states. Election and voting systems. Political suggestions by non-government officials for corrective action, or for changes in existing rules and regulations. Political campaigns and political candidates. Political proposals, laws and actions suggested by non-government officials to be voted on by the entire voting body. The political, social and economic rights to which all creatures are entitled, usually upheld by law. Rights of individuals under civil law. Attempts by any group to control freedoms of speech, religion, and ideas distributed in print, graphics, cyberspace and other ways. Does not include official standards such as cinema ratings, advertising and broadcast standards. Religious rights and freedoms, pressure and intimidation of the believers, censorship in mass media, activities of government bodies, journalistic associations and/or organisations and other NGOs in regards to freedom of belief and practice. Mass media rights and freedoms, pressure and intimidation of the journalists, censorship in mass media, activities of government bodies, journalistic associations and/or organisations and other NGOs in regards to press freedom. Human rights. The whole system for ruling a country. The paid service by civilians for the government, and the often non-paid service of individuals for the benefit of others (public service). People employed by a government at all levels. Individuals, usually elected, who are in public service or commonly in the public eye. Unpaid service for the community as civilian. A written document setting forth the operations of a government and the rights of the citizens therein. Anything involving the protection of one's own country. Armed forces: those employed by a government to conduct war, or to

enforce the security of a nation. Care and provision for those who served in the military. Mandatory service with the national military for a limited period of time. Equipment issued to members of the armed forces. A country's collection of armaments. Means of making a nation, a state, a building or a person secure from harm and outside interference. The philosophy of securing a nation and its borders, as well as the well being of its citizens. Each government department: Divisions of a government that concentrate on specific areas such as health, welfare, economy or war. Heads of state: Symbolic or actual chief representative of a nation, such as royalty, or president or emir, for example. Parliament. Nationalisation: State takeover of private companies or properties. Privatisation: The transfer of state-owned companies or properties to private owners. Safety of citizens. Taxation. Interior policy. Data protection policy. Housing and urban planning policy. Integration policy. Migration of people. Nuclear policy. Regulatory of industry. Food and drink regulations. Cultural policies. International relations. Refugees and internally displaced people. Non-governmental organisation. Political crisis. Disagreement between political groups, usually organised and sometimes resulting in imprisonment of the dissenters. In a few words, we could say that there are nine fundamental sub-categories of the domain Politics:

1. election
 2. fundamental rights
 3. government
 4. government policy
 5. international relations
 6. non-governmental organisations
 7. political crisis
 8. political dissent
 9. political process
- **Religion and belief:** All aspects of human existence involving theology, philosophy, ethics and spirituality. Religious festival or holiday: Holy day or day of observance in a religion which sometimes is a public holiday as well like Christmas, Easter, Ramadan. Religious facilities: Any facility where a group can carry out its religious rites whether it is a mosque, a church, a house or a tent.
 - **Science and Technology:** All aspects pertaining to human understanding of nature and the physical world and the development and application of this knowledge. Biomedical science, mathematics, mechanical engineering, natural science like astronomy and biology, scientific institutions, social sciences like anthropology, archaeology, history, information science, study of law, linguistics, philosophy, political science, psychology and sociology. The rules, national and international, that govern science and technology. These might include international standards setting for equipment and interference standards. Technology and engineering: The study of the practical or industrial

arts, and applied sciences such as physics, hydrodynamics and thermodynamics.

- **Society:** Aspects of the life of humans affecting its relationships. Communities, immigration, denial of services to one or more people because of race, colour, sexual identity, religion, political beliefs etc. Harmful attitudes toward another group based on different ages, racism, religious discrimination, sexism. Family, adoption, courtship, divorce, abortion, marriage, parent and child, mankind, adults, children, disabled, gays and lesbians, gender, infants, national or ethnic minority, senior citizens, teenagers. Social condition like homelessness and poverty. Social problem like abusive behaviour, addiction, juvenile delinquency, prostitution and slavery. Levels of behaviour and its community acceptance. Social issues relating to death and dying. A system of beliefs about acceptable behaviour. Help for those in need of food, housing, health and other services. Philanthropic giving. Extensive health care because of serious illness or disability. Supported programmes to improve the well-being of individuals or groups of individuals.
- **Sport:** Competitive exercise involving physical effort. Organizations and bodies involved in these activities. We could say that there are eight subcategories of this domain:
 1. competition discipline
 2. disciplinary action in sport
 3. drug use in sport
 4. sport event
 5. sport industry
 6. sport organisation
 7. sport venue
 8. transfer
- **Conflicts, war and peace:** Acts of socially or politically motivated protest and/or violence and actions to end them. Especially, act of terror like act of bioterrorism and bombings. Disputes between opposing groups involving the use of weaponry, but not necessarily formally declared a war. Dissatisfaction among the population as evidenced by rallies, strikes, demonstrations or sabotage. The overthrow of an established government by an organised group, often the military or a political party. The death of a large group of people over a brief period of time. Actions to end a war or conflict. Any action to rebuild the society, economy and political system in an area hit by a war. People captured and imprisoned due to social conflicts, political activity, during war or armed conflict.
- **Weather:** The study, reporting and prediction of meteorological phenomena. There are four subcategories:

1. weather forecast: Prediction of the course of the weather in the future either near term or long term
 2. weather phenomena: Special weather conditions
 3. weather statistic: Numerical facts about the weather such as temperature, barometric pressure, river levels, humidity, high and low tides and the like
 4. weather warning: Alerts to the general populace about severe weather coming their way.
- **Other:** this type is selected when the annotator didn't managed to classify the document in one of the previous classes

Metaphor Domains

After the annotation of the document domain, the annotator is searching for metaphors at the certain text. So, the annotator has to decide if the metaphor's domain is:

- Arts, culture & entertainment
- Crime, law & justice
- Disaster & accident
- Economy, business & finance
- Education
- Environment
- Health
- Human interest
- Labour
- Lifestyle & leisure
- Politics
- Religion and belief
- Science & Technology
- Society
- Sport
- Conflicts, war & peace
- Weather
- Other

Having in mind the issues of each document domain, the annotator has to decide the domain of the metaphor. The annotator should select the metaphor domain searching for the right domain of the concrete sentence without taking into account its context. He has just to select the correct domain of the certain phrase. If the annotator considers that the certain metaphor expression does not belong to any of the thematical categories, he can choose the domain "Other".

How to recognize a metaphor?

Metaphor is a figure of speech in which a word or phrase literally denoting one kind of object or idea is used in place of another to suggest a likeness or analogy between them ⁴. At every text have been automatically annotated certain words as possible metaphors (words with red color). The annotators have to verify that the certain words have actually a metaphorical meaning at the concrete context. Also, they have to look for other words with non literal meaning at the certain context which had not been automatically annotated by the system.

In order to find the meaning of a term, the annotators will look up the certain word in the dictionary of Greek Language Manolis Triadafyllidis ⁵. They should have in mind that the basic meaning of a word is more concrete as what it evokes is easier to imagine, see, hear, feel, smell and taste. Moreover, the literal meaning is usually related to bodily action, is more precise and concrete, is historically older and isn't necessarily the most frequent. A lexical unit has to be annotated as metaphorical if its contextual meaning (the meaning observed in a given context) contrasts with its basic meaning, according to the dictionary's definitions, and in order to understand the contextual meaning the annotator has to know the basic meaning of the certain term ⁶. On the contrary, a term is literal if its contextual meaning is similar with the basic definition of the certain term or if it is concrete, precise, or bodily-related enough. At the example "Professional religious education teachers like Marjorie B Clark (Points of View, today) are doing valuable work in many secondary schools" the word valuable is used indirectly in this context because it evokes a referent ("very useful and important") that is different from the basic - direct - meaning of valuable ("worth a lot of money"). Metaphorical meaning arises through non-literal comparison between the contextual and the basic meaning.

Examples of lexical metaphors	Basic meaning	Metaphorical meaning
Valuable work	worth a lot of money	very useful and important
Παροπλισμένοι δημοσιογράφοι	αγκυροβολημένος	Απογύμνωση από αρμοδιότητες, περιορισμός δικαιοδοσίας

In the previous examples, we can see that a term identifies metaphorically a literal term in order to give emphasis on a property of the identified term. In the certain cases, we can say that there are first type's metaphors as just only one term identifies metaphorically another literal term. The literal term is either a single term or a subordinate nominal clause.

In the case of the multi-word metaphorical expressions, the meaning of the expression is totally different of the literal meaning of its constituents. For example, from the constituents of the phrase "My heart swelled with a sea of tears." becomes a new meaning. In the case of the multi-word metaphorical expressions of the type 2, each constituent can be replaced by another term with similar meaning, in contrast with the idiomatic metaphorical expressions of the type 3 which are multi-word

⁴<http://www.merriam-webster.com/dictionary/metaphor>

⁵http://www.greek-language.gr/greekLang/modern_greek/tools/lexica/triantafyllides

⁶It would be useful to be mentioned that the annotator doesn't take into account the historical aspect of the words.

metaphorical expressions but they consist entirely of the specific constituents and no one of them can be replaced by another word with similar meaning. We can say that an idiomatic metaphorical expression is a lexeme made up of a sequence of two or more single lexemes. The lexeme of the multi-word metaphorical expression has properties that are not predictable from the properties of the individual lexemes or their normal mode of combination. Let's see a typical English example of the type 3 of metaphors which is "kick the bucket". This phrase means to die rather than to hit a bucket with one's foot. In the case of Greek language a typical example of idiomatic metaphorical expression may be considered the phrase «κρούω τον κώδωνα του κινδύνου». This phrase means *εφιστώ την προσοχή κάποιου σε ορισμένο θέμα* rather than *χτυπώ το κουδούνι του κινδύνου*. We should refer that the certain phrases have a typical structure and it is very rare to appear another structure or their lexemes have predicates. The components of the multi-word metaphorical expression must obey the following functions:

- Substitutability: It cannot be replaced a component of the certain expression by a synonymous, hyperonymos or a a synoonymous word.
- Deletion: It cannot be deleted a component of the multi-word metaphorical expression.
- Category transformation: A component of the multi-word expression cannot change lexical category.
- Permutation: It is not possible to change the position of a term within the multi-word expression.
- Semantic Opacity: The sequence of the words cannot be understood, for example *τα έκανε γης Μαδιάμ* = *προκάλεσε μεγάλη καταστροφή, αναστάτωση*. Opaque expressions are derived from hystorical or mythological events, from religious and literary tradition or referred to earlier habits and conditions that are currently ignored by the midfielder speaker of a language ⁷.
- Translatability: As the multi-word metaphorical expressions are original creation of each language, the translation of these expressions to another language cannot be a word to word literal translation.

In the case of the lexical metaphor and the IS-A metaphor only a term has a non literal meaning and enhances other terms with properties which constitute typical characteristics of the metaphorically used term. Regarding to lexical metaphors and IS-A metaphors the metaphorical term cannot be replaced by another term with similar meaning as this metaphor links the properties of the certain terms and changing the properties of the terms is plausible to change the metaphor, too. We can examine two examples; «Ο Γιάννης είναι γαϊδούρι» και «Ο Γιάννης είναι αλεπού» although the donkey and the fox are animals each of them gives a different property to the subject. Especially, in the first case the subject is unconcerned while in the second the same subject is very intelligent.

⁷Αναστασιάδη-Συμεωνίδη & Ευθυμίου, 2006: 46

	Literal meaning	Non literal example
Έφαγα τον κόσμο	eat	Search everywhere
Το μωρό έφαγε γάλα	eat	eat
Το μωρό κατανάλωσε γάλα	eat	eat
Το μωρό έλαβε γάλα	eat	eat
*Κατανάλωσα τον κόσμο	*eat	*eat
*Έλαβα τον κόσμο	*eat	*eat

We should refer that in the case of the direct, IS-A metaphor may be not a contrast between the contextual and a more basic meaning. The contextual meaning is also and the basic meaning. The comparison is expressed through direct language use. The direct language use may or may not be signalled with words like “like” or «σαν» in the case of Greek language. It is mentioned an example (from Amsterdam Metaphor Corpus) of a direct, IS-A metaphor without a signal:

“They [system developers] seem to think that you can ask a businessman what his requirements are and get an answer that amounts to a draft system specification. A doctor doesn’t ask his patient what treatment to prescribe. The patient can explain only what the problem is. It is the doctor that provides the remedy. A user may have a deep knowledge of business problems, but knowing little about computers, has no idea how they should be tackled.”

The lexical units in brown have been marked as direct metaphor. Their direct metaphorical use is not signalled. Nevertheless, there is a comparison between two different domains (medical and systems development).

Moreover, within the context of lexical metaphor or multi-word metaphorical expression may have impersonation of objects or concepts and conversely animate terms may be used as inanimate, for instance «Η σκέψη του με κυρίευσε» where is obvious the impersonation of the concept of thought. Respectively, at the example «γκρέμισαν τον δικτάτορα» the concept of dictator is used as an object which may be demolished. The metaphorical term of the phrase has not a semantic connection with the usual arguments of the literal term. Looking at a the example «έφαγα τον κόσμο να σε βρω», we can ascertain that the term “word” differs semantically from the typical arguments of the verb “eat” which may be foods, fruits and others. So, a native speaker knows the typical literal arguments of a word. In order to find out the metaphorical usage of a term is sufficient to try to replace the certain term with another term which is known that has a literal meaning at this context. The successful replacement indicates the literal usage of the term otherwise it is obvious that the term has a metaphorical function at the certain context.

We can observe that although the terms may be synonymous, when one of them has a metaphorical meaning, it cannot be replaced by another synonym term in the same context.

At the same time, we can try to check if the term which seems to be used metaphorically can take the same modifiers with a term which is known that has a literal usage at this context. Let’s consider the previous example. A typical literal

instance may be «Έφαγα ένα καλοψημένο κοτόπουλο». Now, we can check if the modifier “well-done” can determine and the word “world” in order to investigate if this term has the same usage with the already known example.

Example	Modifier	Meaning
Έφαγα τον κόσμο	Without modifier	Non literal
Έφαγα το κοτόπουλο	Without modifier	Literal
Έφαγα το καλοψημένο κοτόπουλο	With modifier	Literal
*Έφαγα τον καλοψημένο κόσμο	With modifier	

Looking at the previous examples, we can observe that when two terms have different meanings, they can take the same modifiers, even if they appear at the same context. So, in order to investigate the meaning of a term, we could think about another literal term which may appear at the same context. If the two terms may receive the same modifiers, they have and the same meaning. Otherwise, one of them has a non literal meaning at the certain context.

Simultaneously, the syntactic function of the metaphorical term may differ from the typical syntactic function of this term. At the example «Ο άντρας πολιορκεί το οχυρό» we can understand that a man overpowers fortified position, while at the example «Η γυναίκα πολιορκεί τον άνδρα» we can see the different syntactic function of the term “man”, which is the object of the verb and accepts the woman’s actions like being a fortress.

Between the two cases, the word “man” has a different syntactic function. In the first case, the man is the person who acts while in the second case is the person who receives the result of the action. Looking up the meaning of the verb in dictionary, we can realize that the two objects have a different meaning as the word “man” and the word “stronghold” cannot have the similar meaning at the same context.

We should have in mind that the cases of indirect, lexical metaphors, multi-word metaphorical expressions and direct, IS-A metaphors between their constituents may appear adjectives, adverbial determinations or even prepositional determinations in order to accentuate the meaning of the metaphorical expression. In this case, the annotator has to select as metaphorical the whole part of the expression included the determinations.

Metaphor detection using Ellogon

In order to recognize if a phrase has a literal or a non literal meaning at a certain text, we are going to use tool Ellogon⁸. Ellogon is a multi-lingual, cross-platform, general-purpose language engineering environment, developed in order to aid both researchers who are doing research in computational linguistics, as well as companies who produce and deliver language engineering systems. The fundamental documentation of the tool are available on the link of the footnote. In order to make the adjustment of the annotators with the tool easier, the annotators will be trained and supported over the tool.

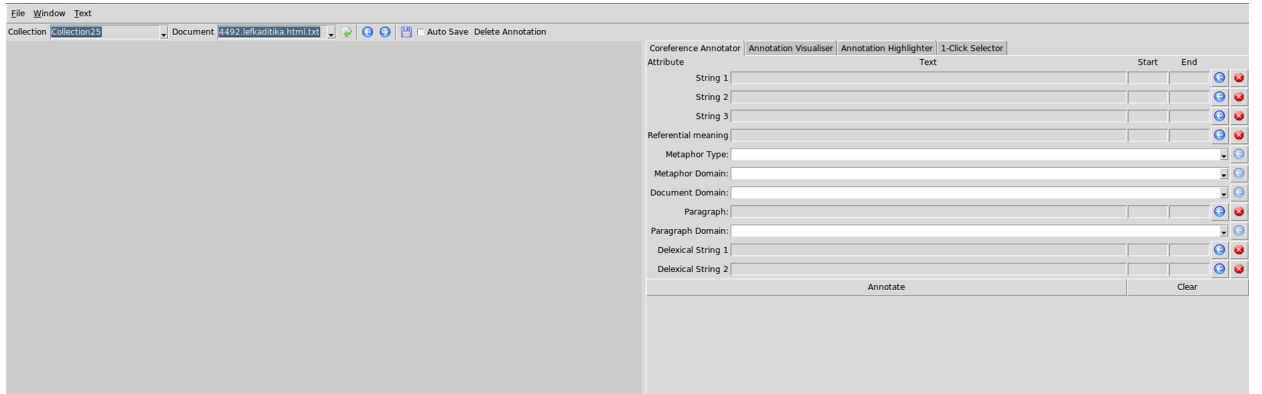


Figure 1: Screenshot for the environment of tool

⁸<http://ellogon.org/>

Steps of Metaphor Identification Procedure

1. Read carefully the whole text in order to get a general understanding of it
2. Classify the text in one of the eighteen document domains
3. Determine the lexical units considering whether the words, that had already been annotated as metaphors, have actually a metaphorical meaning at the certain context and searching for other metaphorical words which had not been annotated by the system
4. Establish the contextual meaning of each lexical unit
5. Determine if it has another more basic (concrete, body-related, precise, historically older) meaning than the contextual meaning. This basic meaning is not necessarily the most frequent meaning
6. If the contextual meaning contrasts with the basic meaning and it can be understood in comparison with it then the lexical unit should be marked as metaphorical
7. Classify the metaphor in one of the eighteen metaphor domains
8. Searching for metaphors, if you detect delexical verbs, annotate them

The screenshot displays a software interface for analyzing text. On the left, a document titled 'Collection: 4492.lefkadika.html.txt' is open, showing a Greek article about the 'Pasi' (Pasi) movement. The text is annotated with various tags and highlights. On the right, a table titled 'Metaphor Identification' lists 24 identified metaphors. The table has columns for ID, Type, Spans, and Attributes. The attributes column contains detailed information about each metaphor, including its type (e.g., 'type=dm:Health', 'type=dm:Science and Technology'), its modified by user (e.g., 'modified_by_user=Marianna@143.233.5.6'), and its type (e.g., 'type=dm:Health', 'type=dm:Science and Technology').

ID	Type	Spans	Attributes
0	metaphor	0:6	type=dm:Health, modified_by_user=Marianna@143.233.5.6
1	metaphor	358:380	type=dm:Health, modified_by_user=Marianna@143.233.5.6
2	metaphor	358:380	type=dm:Health, modified_by_user=Marianna@143.233.5.6
3	metaphor	431:456	type=dm:Health, modified_by_user=Marianna@143.233.5.6
5	metaphor	431:456	type=dm:Health, modified_by_user=Marianna@143.233.5.6
6	metaphor	744:776	type=dm:Health, modified_by_user=Marianna@143.233.5.6
7	metaphor	752:761	type=dm:Health, modified_by_user=Marianna@143.233.5.6
8	metaphor	744:776	type=dm:Science and Technology, modified_by_user=Marianna@143.233.5.6
9	metaphor	752:761	type=dm:Science and Technology, modified_by_user=Marianna@143.233.5.6
10	metaphor	752:776	type=dm:Health, modified_by_user=Marianna@143.233.5.6
11	metaphor	752:776	type=dm:Health, modified_by_user=Marianna@143.233.5.6
12	metaphor	778:835	type=dm:Health, modified_by_user=Marianna@143.233.5.6
13	metaphor	778:835	type=dm:Health, modified_by_user=Marianna@143.233.5.6
14	metaphor	837:898	type=dm:Health, modified_by_user=Marianna@143.233.5.6
15	metaphor	837:898	type=dm:Health, modified_by_user=Marianna@143.233.5.6
16	metaphor	1021:1044	type=dm:Health, modified_by_user=Marianna@143.233.5.6
17	metaphor	1021:1044	type=dm:Health, modified_by_user=Marianna@143.233.5.6
18	metaphor	1046:1066	type=dm:Health, modified_by_user=Marianna@143.233.5.6
19	metaphor	1046:1066	type=dm:Health, modified_by_user=Marianna@143.233.5.6
20	metaphor	1452:1472	type=dm:Health, modified_by_user=Marianna@143.233.5.6
21	metaphor	1452:1472	type=dm:Health, modified_by_user=Marianna@143.233.5.6
22	metaphor	1569:1598	type=dm:Health, modified_by_user=Marianna@143.233.5.6
23	metaphor	1569:1598	type=dm:Health, modified_by_user=Marianna@143.233.5.6
24	metaphor	2097:2122	type=dm:Health, modified_by_user=Marianna@143.233.5.6

Figure 2: Screenshot from an annotated article

Example

Let's search for metaphorical expressions at the following paragraph.

“Η ελληνική κυβέρνηση επέλεξε τον ρόλο του **σιωπηλού παρατηρητή** που κατάφερε μέσα σε λίγες μέρες να προκαλέσει **ανυπολόγιστες** καταστροφές στην κυπριακή οικονομία και ταυτόχρονα να αυξήσει δραματικά τον **συστημικό** κίνδυνο για ολόκληρη την **ευρωζώνη**. Για την Κύπρο οι εξελίξεις θα είναι δραματικές. Η απότομη συρρίκνωση του τομέα των χρηματοπιστωτικών...”

If we had read the whole text, it would be easy to ascertain that it is a part from a text which deals with political matters. So, the document domain of this text will be “Politics”. As we can see, there are certain words which had been annotated as possible metaphors. So, the annotators have to verify that these words have actually a metaphorical meaning at the certain context and also they have to look for other metaphorical words which had not been annotated by the system but they have a metaphorical meaning at the certain text.

The phrase «τον ρόλο του **σιωπηλού παρατηρητή**» has a non literal usage at the certain context. This metaphor is type 1. The domain of the certain metaphor is the “Arts, culture & entertainment”.

At the same sentence, the system has annotated the word «**ανυπολόγιστες**» which has a metaphorical meaning with the rest context «καταστροφές στην κυπριακή οικονομία». The whole phrase is a first type metaphor and belongs to the metaphor domain “Disaster and Accident”.

The phrase «αυξήσει δραματικά τον **συστημικό** κίνδυνο» is metaphorical and not only the word «**συστημικό**» which just operates as an indicator of a metaphorical phrase. At the same phrase the word «**ευρωζώνη**» has been annotated as a metaphorical indicator. However, the certain word at the concrete context has a literal meaning. So, the annotator, shouldn't consider this word as an indicator of a possible metaphor.

At the same text there is the phrase «Για την Κύπρο οι εξελίξεις θα είναι δραματικές.». As we can see at this phrase there is not any metaphorical indicator. However, the phrase «οι εξελίξεις θα είναι δραματικές.» is a fourth type metaphor which belongs to the metaphor domain “Arts, culture & entertainment”. For this reason, the annotator shouldn't bear in mind only the annotated words but he must search for other words which hadn't been annotated as metaphorical indicators. The system may have not annotated even typical metaphorical words as the word «φάυλος» at the phrase «φάυλος κύκλος», which appears at the concrete text and this phrase is a characteristic non literal phrase of type 3. In this case, as we have already told, the annotator hasn't to refer the metaphor domain.

Common Misunderstandings

1. We cannot annotate as a metaphor just only a word as a metaphor is a figure of speech in which a comparison is made between two unlike things that actually have something in common. So, it is necessary to annotate both of the unlike things in order to be found out the common property.
2. If we annotate a metaphor as idiomatic metaphorical expression (type 3), we don't annotate the domain of the certain metaphor as this expression consists entirely of the specific constituents.
3. We should have in mind that we concentrate on detecting metaphorical expressions. For this reason, we don't annotate either metonymy or personalization.
4. In a special category we annotate the delexical verbs, which must be distinguished from the metaphorical expressions of the type 3. In the case of the delexical verbs the verb gives only grammatical information and the noun declares the action or the state. In the case of the metaphorical expressions of the type 3, a new meaning becomes due to all the constituents of the phrase.
5. The metaphorical expressions of type 3 are characterized by lexicogrammatical fixedness, while the metaphorical expressions of type 2 have a lexicogrammatical and syntactical variety and their constituents can be replaced by another term.
6. If there are determinations, we annotate the metaphorical expression including them as they accentuate the meaning of the metaphor.
7. The metaphors of type 1 are not composed exclusively of only two words. The metaphorical expression may include a term that functions non literally and a subordinate nominal clause.
8. The paragraph domain should be annotated if and only if a paragraph belongs to different domain from the domain of document.

Appendix A

Idiomatic metaphorical expressions

1. αυτοί...στην κοσμάρα τους
2. κάθεται στα αβγά του
3. τα έβαψα μαύρα
4. την πάτησα
5. παίζω στα δάχτυλα
6. μένω με σταυρωμένα χέρια
7. ρίχνω μαύρη πέτρα
8. μαυρίζω στο ξύλο
9. βγάζω από τη μέση
10. λέω μεγάλα λόγια
11. τα έκανε θάλασσα
12. τα έκανε λίμπα
13. έδωσαν τα χέρια
14. δίνε του
15. έχει μεσάνυχτα
16. παίρνω στροφές
17. παίρνει ανάποδες στροφές
18. του έριξε το γάντι
19. δεν το βάζει κάτω
20. μας την έφερε
21. την πάτησα

The previous phrases are typical examples of idiomatic metaphorical expressions. The certain phrases may maintain the concrete structure but may have their verb in another grammatical person or in another tense. However, their meaning remains unaltered.

Appendix B

Copula

As copula we can define a connecting word, in particular a form of the verb be connecting a subject and complement. Apart from verb be, in greek language there are and other copulas, characteristic cases of them may consider the verbs γίνομαι, δείχνω, φαινομαι, εμφανίζομαι, παρουσιάζομαι, αποδεικνύομαι, φαντάζω, λέγομαι, κατάντώ, εκλέγομαι, διορίζομαι, υπήρξα, στάθηκα, χρημάτισα, διετέλεσα, μοιάζω, μένω, παραμένω, απαρτίζω, αποτελώ, συγκροτώ, συνθέτω, συνιστώ, στοιχειοθετώ