The Construction of Spatial Invasion of Carcinogen Risk in Scientific Discourse: A Corpus- Based Study

Huda H. Khalil [a]
huda.hadi@coeduw.uobaghdad.edu.iq
University of Baghdad, Iraq

Nassier A. G. Al-Zubaidi [b]
naghubin@coart.uobaghdad.edu.iq
University of Baghdad, Iraq

ABSTRACT

Scientific discourse is devoted, in many occasions, to transmitting knowledge about the risk of proximity to carcinogens; e.g. living near nuclear power station may cause cancer because of the emitted radiation which spreads around. No previous study has attempted to investigate the linguistic manifestation of the spatial perspective of carcinogen risk. The present paper aims at investigating the construction of the spatial invasion of carcinogen risk in scientific discourse to promote people to take preventive measures. To achieve this aim, Cap's (2013) proximization theory of crisis and treat construction is employed. The theory provides three proximization strategies: spatial, temporal and axiological. The spatial proximization strategy, in particular, is adopted in the analysis. The analysis procedure is both qualitative and quantitative. Calculations are performed by corpus linguistics. AntConc software is used for this purpose. The corpus analysis tools used are word list, concordance, file view, cluster/ N- gram, wildcard * and the file view tool. The corpus consists of a set of scientific articles which are combined by the researchers to form the corpus. Results have revealed that scientific discourse employs various linguistic tools to construct the special proximization of carcinogen risk. However, the linguistic tools are employed with different rates to achieve certain cognitive pragmatic aims.

Keywords: carcinogen; prevention; proximization theory; scientific discourse; spatial proximization

INTRODUCTION

Many scientific reports emphasize the increasing rates of cancer cases and mortality. "About one-fifth of people worldwide and one-third of people in industrialized countries will be diagnosed with cancer during their lifetimes" (Gatto, 2021, p. 1). According to World Health Organization (WHO), 20% of cancers worldwide are estimated to be attributed to environmental risks such as management of chemicals, air pollution, and radiation and UV rays from sunlight. Cancer cases that are caused by environmental agents vary from one country to the other depending on the degree of industrialization (Gatto, 2021, p. 1). Scientists mostly accuse environmental carcinogens as a major spatial factor in causing cancer. They assure that environmental prevention can contribute to reduction in cancer potential. Environmental carcinogen factors include air pollution,

[a] *Main author*
[b] *Corresponding author*

eISSN: 2550-2131
ISSN: 1675-8021

exposure to occupational carcinogen compounds, tobacco smoke, certain lifestyle factors, processed food, etc. (Denholm, Schüz, Straif, Ali, Bonas, Gjebrea, Sifton & Olsson, 2016, p. 220).

A carcinogen is an agent or substance that is of cancer-causing potential in humans (Carcinogen, 2008). Cancer and carcinogens are the most important topics that scientific discourse has laid heavy emphasis on. Scientists from different fields find themselves committed to furnish the world with information about cancer. Therefore, scientific discourse has always been a rich source of such information.

Although exposure duration to carcinogens and the amount of their consumption are vital factors for getting cancer, scientists mainly focus on the space as an essential domain in which carcinogens play essential role as physical entities. Many scientists rely on spatial factors for the transmission of carcinogens from the environment to human bodies. Hence, the present paper aims at investigating the construction of spatial invasion of carcinogen risk in scientific discourse. It attempts to answer the question: What are the linguistic tools (with their rates of utilization) that the scientific discourse employ to construct the special invasion of carcinogen risk? Since environmental carcinogens move from one place to the other in our planet, the spatial invasion represents an important aspect of their risk. Distance reduction (proximization) between people and carcinogens is the essence of spatial invasion. To achieve this aim, Cap's (2013) theory of proximization is employed for analysis. The theory consists of three proximization strategies: spatial, temporal and axiological. Of these three strategies, the spatial proximization (SP) strategy is adopted in the analysis of the data.

SCIENTIFIC DISCOURSE

Discourse is the result of social practice. People create a specialized discourse as a strategy for coping effectively with certain special global events (Gustilo, May Pura & Biermeier, 2021; Younus, 2020, p. 14). Such theoretical approach emphasizes the shared and conventional social ways of interacting, behaving and communicating (Hanauer, 2006, p. 3). There is also the idea of discourse as individual cognitive development. The basic principle of this idea is the "interaction between the wider experiences in the social context and the internal cognitive system of the individual" (Aajami, 2020, p. 3; Hanauer, 2006, p. 4). Scientists and scientific discourse producers construct knowledge through experiments, laboratory tests, field work, surveys, objective observation, etc. which all occur within a wide social context. Therefore, contextualized knowledge is textualized later on in scientific discourse where scientific concepts (knowledge) are constructed through language (Idris & Ghani, 2012, p. 69).

Discourse studies are characterized by multidisciplinarity because they go beyond the linguistic domain to various cognitive, literary, psychological, social, anthropological and other domains (van Dijk, 2002). Discourse is a contextualized activity that functions in certain ways according to certain contexts (Hanauer, 2006, p. 2). Thus, discourse participants do not limit themselves to their knowledge of the language system. Rather, they rely on the knowledge of the context in which the communicative acts take place. It is this latter knowledge which guides the use of utterances and dictates their structure (Adegbite, 2000, p. 63; Johnstone, 2008, p. 10).

In scientific discourse, the packaging of empirical data might be the main work for researchers. However, the data, by itself, is not self-explanatory. Rather, "it is the discoursal construct within which the data is conceived that makes it meaningful to the development of scientific concepts" (Hanauer, 2006, p. 21). What is central within the scientific discovery activities is scientific thinking and scientific discourse. This involves the admission of science as

a field in which a theory is socially constructed and that scientific knowledge depends on language which provides a tool for the expression of hypotheses, scientific theories and conceptualizations (Goh, Chan, Ali & Rashid, 2019, p. 221; Hanauer, 2006, p. 21). When texts are produced for readers, the presented knowledge causes certain modification in the cognitive system of readers who are now presented with information that might change their attitude and behavior. Therefore, discourse is more related to language in use in the sense of where language is socially situated (Paltridge, 2000, p. 3; Yu, 2009, p. 50).

A number of previous studies have tackled scientific discourse from different perspectives. Myers (2003) has challenged the "mainstream view" of science popularization (which holds that it is a one-way simplifying process) with scientific papers serving as the original sources of knowledge that are then debased by translation for an audience that is ignorant of such issues. Overton (2013) has investigated the usage of certain words and their effects on the philosophy of science. Wei and Yu (2019) have discussed some of the formal features of scientific discourse and their effects on translating such kind of discourse. Khalil and Al- Zubaidi (2022a) have examined the way in which ideological conflict in scientific discourse leads to constructing carcinogen risk. Abbas (2020) has shown the types of overlap that overwhelm scientific discourse of therapy sessions and the way the overlap is managed and resolved. Moreover, Khalil and Al-Zubaidi (2022b) have conducted a contrastive study to compare between temporal proximization in both English and Arabic scientific discourse. The present paper tries to fill a gap in literature. It tries to respond to inquiries about the way in which carcinogen risk is linguistically constructed through spatial related concepts in English scientific discourse.

PROXIMIZATION THEORY

Proximization theory is a cognitive pragmatic theory. P*roximization* represents a rather recent notion in linguistics. It was first proposed by Cap (2005) who defines it as "an organized, strategic deployment of cognitive-pragmatic construals of/ in (originally, political) discourse" (Cap, 2013, p. 5). Proximization "has developed into a cognitive-linguistic, pragmatic, as well as a critical discourse analytic concept which accounts for the symbolic construal of relations between entities within the Discourse Space (DS)" (Chilton, 2005 cited in Cap, 2013, p. 5). Proximization is concerned with the symbolic shift of the deictic elements from the deictic periphery to deictic center within the DS. Referring to proximization, Cap (2018, p. 97) states:

> [P]roximization is a discursive strategy of presenting physically and temporally distant events and states of affairs (including 'distant' adversarial ideologies) as increasingly and negatively consequential to the speaker and her addressee. Projecting the distant entities as gradually encroaching upon the speaker-addressee territory (both physical and ideological), the speaker seeks legitimization of actions and/or policies she proposes to neutralize the growing impact of the negative, 'foreign', 'alien', 'antagonistic', entities.

Cap (2020, p. 281) identifies proximization as a discursive strategy for the construction of conflict, crisis and threat which is achieved by "the movement dynamics of entities positioned in Discourse Space". Proximization allows the physical and temporal presentation of distant states of affairs and events in relation to the interlocutors in the deictic center of the DS. Proximization has been mostly used in political discourse studies to investigate the rhetoric of war (Okulska & Cap, 2010), the construction of anti-migration discourse (Hart, 2010), designing foreign policy

documents (Dunmire, 2011), political party representation (Kaal, 2012) and the construction of national memory (Filardo Llamas, 2010).

Proximization theory employs proximization as a force construal process to achieve "closeness of the external threat in order to solicit legitimization of preventive measures" (Cap, 2020, p. 281). Prevention represents the preventive measures which are globally implemented to control affliction. Risk is the reason behind the execution of preventive measures. Proximization theory puts forward a set of lexico-grammatical choices which establish both the deictic center and the deictic periphery to enhance symbolic construals of the peripheral entities as they cross the distance in DS to settle in the deictic center. These choices are obtained from the cognitive classes of space, time and value to build up spatio-temporal-axiological (STA) proximization model (Cap, 2013, p. 7- 9).

Threat is initiated by DS-peripheral entities which are considered as outside-deictic-center (ODCs). The ODCs cross the Space to invade the inside-deictic-center (IDC) entities. ODCs are represented with negative attitudes since they are threatening and harmful to the IDCs which are represented with positive attitudes. Interlocutors are usually part of the IDCs. The negative representation of the invading ODCs raises fear and promotes preventive measures. Thus, the target is to enhance the public approval in order to motivate preventive action (Cap, 2020, p. 281).

Since threat is of a spatio-temporal and ideological nature, proximization can be of three aspects. First, it is SP which is a forced construal of the DS peripheral entities that physically invade the DS central entities. SP is formed of conceptualizations in spatial terms. Second, it is temporal proximization which is a forced construal of the imminent, momentous and historic conflict which requires instant response and preventive actions (Cap, 2020, p. 281). Third, it is the axiological proximization that "involves construal of a gathering ideological clash between the home values of the DS central entities (IDCs) and the alien and antagonistic (ODC) values" (Cap, 2020, p. 282). ODC values "are construed to reveal potential to materialize (… prompt a physical impact) within the IDC … home territory" (Cap, 2020, p. 282). Proximization is presented in Figure (1):

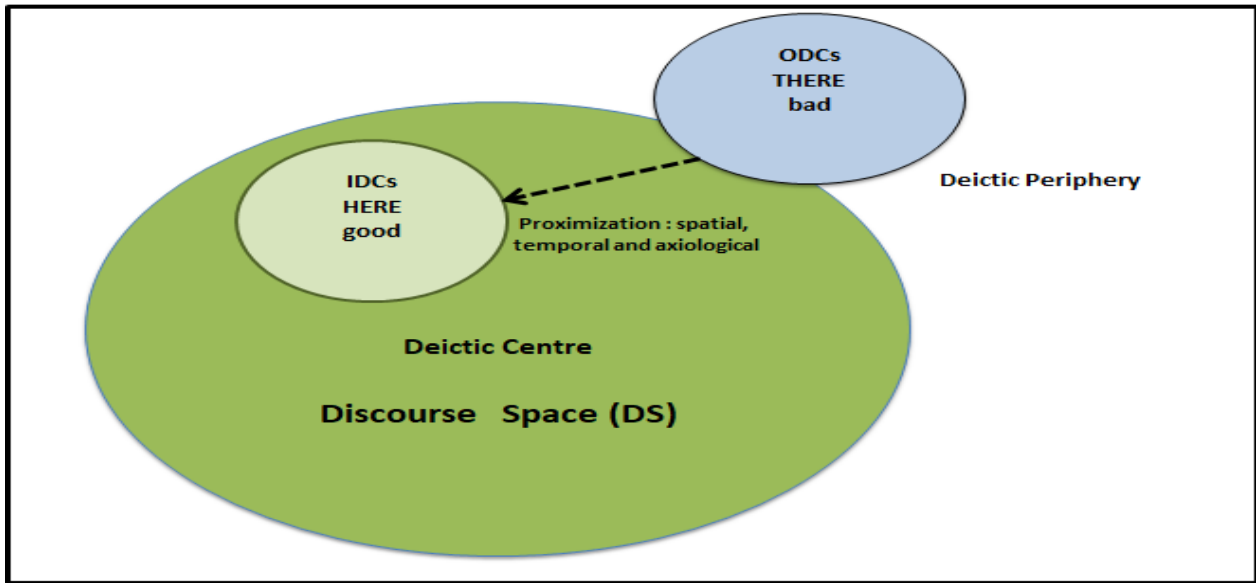

FIGURE 1. Proximization framework (Cap, 2020, p. 282)

SPATIAL PROXIMIZATION

SP is "a forced construal of the Discourse Space (DS) peripheral entities (ODCs) encroaching physically upon the DS central entities (IDCs) located in the deictic center of the Space" (Cap, 2013, p. 105). It imposes the view of a destructive character for the ODCs' influence. The central entity to that view is considered to be the threat that is construed in physical terms. SP postulates the existence of two kinds of entities in the DS: central and peripheral. They are both construed in physical terms. Therefore, two categories are needed to accommodate the linguistic items that denote these entities. Cap (2013, p. 105) defines these categories as follows:

> "(1) Noun phrases (NPs) construed as elements of the deictic center of the DS (IDCs)
> (2) Noun phrases (NPs) construed as elements outside the deictic center of the DS (ODCs)"

In both categories, many items are single nominal words (e.g. chemicals, women) or nominal phrases that are compacted syntactically (e.g. carcinogen risk, college students) "denoting the IDCs and ODCs as physical entities, as well as embodiments and carriers of positive (IDC) and negative (ODC) values" Cap (2013, p. 105). Consequently, ODCs are portrayed as direct and strong threat to the IDCs, and category 2 is assigned for the threatening elements. Since the ODC entities in category 2 invade the deictic center where the IDC entities in category 1 exist, a category that accommodates "markers of the symbolic movement between the DS periphery and the center" is needed. Such markers must include verbal forms. So, category 3 is postulated as follows:

> "(3) Verb phrases (VPs) of motion and directionality construed as markers of movement of ODCs towards the deictic center"
>
> (Cap, 2013, p. 106)

The first three categories represent 'the "canonical" structure of SP: the markers of DS-central ("home") and DS-peripheral ("foreign") entities are all there, as are the markers of negative impact of the foreign entities upon the home entities'(Cap, 2013, p. 107) . Although SP imposes construals of destructive physical impact, the construals vary depending on "how far" (the distance) 'the destructive scenario goes; whether it includes "just" the anticipation and then the invasion part or perhaps also the vision of the effects/consequences' (Cap, 2013, p. 107). Therefore, Cap (2013, p. 107) has added "extra" three categories as follows:

> (4) Verb phrases (VPs) of action construed as markers of impact of ODCs upon IDCs
> (5) Noun phrases (NPs) denoting abstract concepts construed as anticipations of impact of ODCs upon IDCs
> (6) Noun phrases (NPs) denoting abstract concepts construed as effects of impact of ODCs upon IDCs

Categories 4, 5 and 6 may comprise lexical items like *destroy*, *threat*, and *catastrophe*, respectively. The lexico- grammatical tools of category 4 comprise action verbs. While category 4 portrays destruction as an impact of the ODCs upon the IDCs, category 5 represents the anticipated threat (as an abstract concept) of the impact of the ODCs upon the IDCs. Category 6 represents the catastrophe (as an abstract concept) which comes as an effect of the impact of ODCs upon IDCs. The six categories that constitute the SP strategy are put in table (1):

eISSN: 2550-2131
ISSN: 1675-8021

TABLE 1. SP strategy

Category	Lexico- grammatical manifestation within the discourse space
1	Noun phrases (NPs) construed as elements of the deictic center of the DS (IDCs)
2	Noun phrases (NPs) construed as elements outside the deictic center of the DS (ODCs)
3	Verb phrases (VPs) of motion and directionality construed as markers of movement of ODCs towards the deictic center
4	Verb phrases (VPs) of action construed as markers of impact of ODCs upon IDCs
5	Noun phrases (NPs) denoting abstract concepts construed as anticipations of impact of ODCs upon IDCs
6	Noun phrases (NPs) denoting abstract concepts construed as effects of impact of ODCs upon IDCs

METHODOLOGY

To achieve the aim of the present paper, Cap's (2013) proximization theory is adopted for analysis. The theory can be applied to carcinogen risk discourse where carcinogens are metaphorically construed as enemies that threaten interlocutors. Cap (2017, p. 35) has extended the applicability of the theory to investigate the construction of fear in health and disease prevention discourse depending on "three interrelated threat construals: the construal of the ODC as an enemy entity, the construal of the ODC impact speed and the construal of the ODC impact consequences". The ultimate aim is to wage a preventive war against carcinogens. The war metaphor has always been used to describe and fight cancer since '1971, when US president Richard M. Nixon declared a federal ''war on cancer'' with the national cancer act' (Cap, 2014, p. 23). In addition, in 1978, Susan Sontag has published her seminal book *Illness as Metaphor* which enhanced the public awareness of diseases within a metaphorical sense. Enriching the public awareness on disasters (such as carcinogen invasion) is an epistemological process for empowering the knowledge of people on the consequences of these disasters to create a more "self-reliant community" (Selvaraj & Sandaran, 2019, p. 124).

The metaphoric correspondence between war and carcinogen risk in the present paper is implemented in this way: there is an enemy (a carcinogen), a commander (the scientist who warns people about carcinogens), a combatant (the actor/ text producer), allies (medical teams) and weapons (preventative measurements). When the risk of having cancer is located within the DS, carcinogens are the peripheral entities (ODCs), and human bodies are the central entities (IDCs) as illustrated in figure 2:

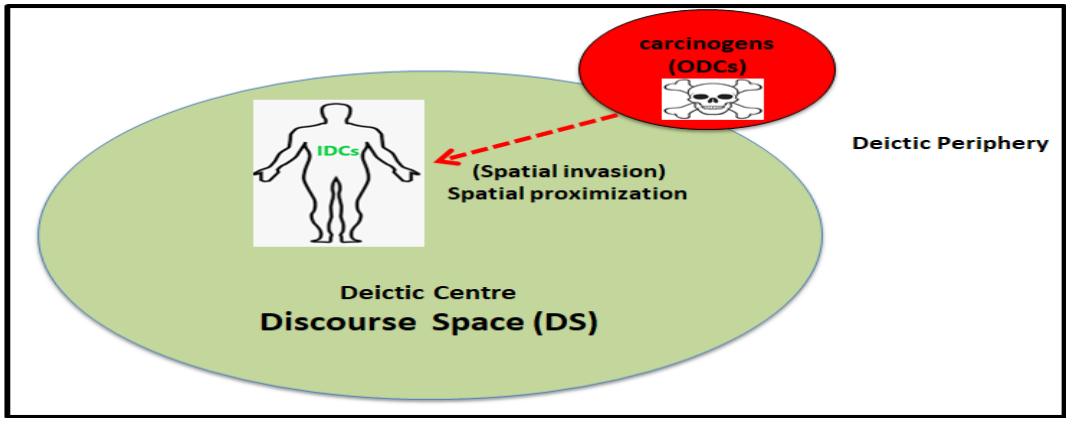

FIGURE 2. Carcinogen spatial invasion (Adopted from Cap, 2020, p. 282)

SP is the strategy adopted for analysis. The analysis procedure is both qualitative and quantitative. Calculations are performed by corpus linguistics using Anthony's AntConc (2019) (version 3.5.8) corpus linguistics software which produces advanced options and tools which are freely available online (Tabbert, 2016, pp. 56- 57). AntConc provides basic text analysis features, and it is capable of working with uploaded corpora. It is user-friendly, straightforward and lightweight program that is easy to use as corpus analysis toolkit (Diniz, 2005, p. 26; Jaafar & Jasim, 2022).

To build the corpus, texts have been extracted from electronic (online) sources and converted to Word files (a file for each text/ article). AntFileConverter has been used to convert the word files to txt. format to be processed by AntConc. Figure 3 illustrates the general framework of the methodology:

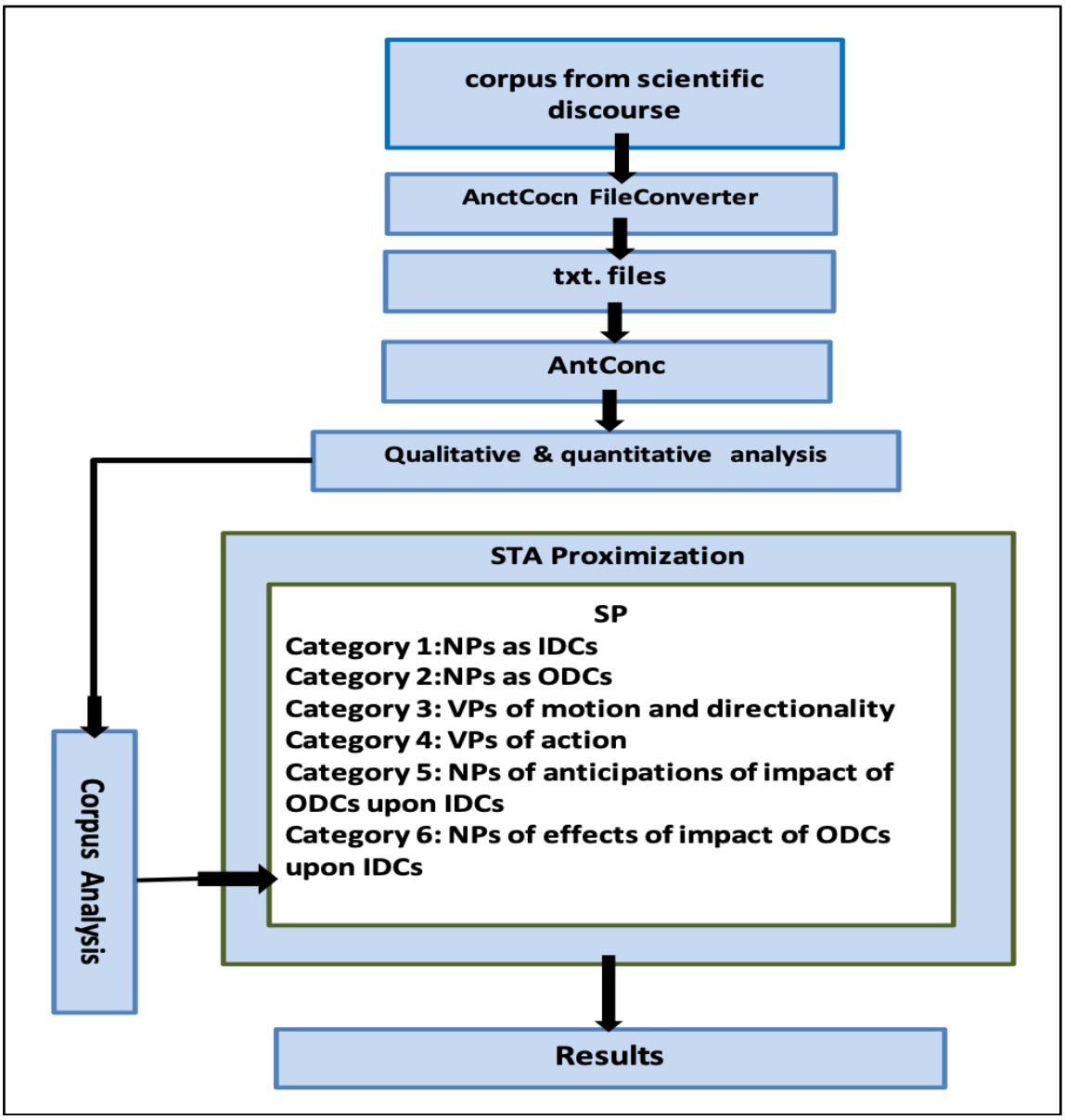

FIGURE 3. Methodology

DATA DESCRIPTION

The paper targets scientific discourse where scientists, researchers and individuals of knowledge from different scientific fields express their concerns of environmental carcinogens. The corpus was built by the researchers who collected the texts from different sources to form the corpus. Preparing the corpus is based on the criteria of type, genre, source and size. The type of discourse targeted is scientific discourse which can be realized through genres like research papers, dissertations, monographs, news reports on scientific facts, conference proceedings, theses, scientific reports, posters, review articles, summaries, abstracts, etc. (Mordovina & Nikulshina, 2010, p. 140). As for the size of the corpora, the evaluation of corpus size can be estimated depending on either the number of documents (articles) or the total number of tokens (words) (Lewis, 2001; cited in Skier & Vibulphol, 2016, p. 241). There is no agreed upon limit to the number of tokens in a corpus since this issue is highly dependent upon the research questions and the linguistic features under investigation (Brezina, 2018, p. 18). Corpus users must critically think "about the nature of the evidence that the corpora provide in terms of their quality (representativeness and balance) as well as their quantity (corpus size)" (Brezina, 2018, p. 19). Thus, a representative corpus may consist of thousands, millions or even billions of words (Brezina, 2018, p. 38).

The paper attempts to find out how scientific discourse employs specific linguistic tools to construct the spatial invasion of carcinogen risk in people's life. The corpus has been combined and set up by the researchers; it is not readily provided by a corpus bank or any other web source that provides million or billion word corpora collected from vast number of genres. The kind of discourse that the paper targets is specific (scientific) and the topic is also specific (carcinogens). In addition, the corpus consists of scientific articles from different genres and the number of articles is a subjective measurement for the size of the corpora. Therefore, the number of tokens is adopted for this purpose as a measurement standard for the size of the corpus and the corpus consists of 56410 tokens as is shown in figures (4):

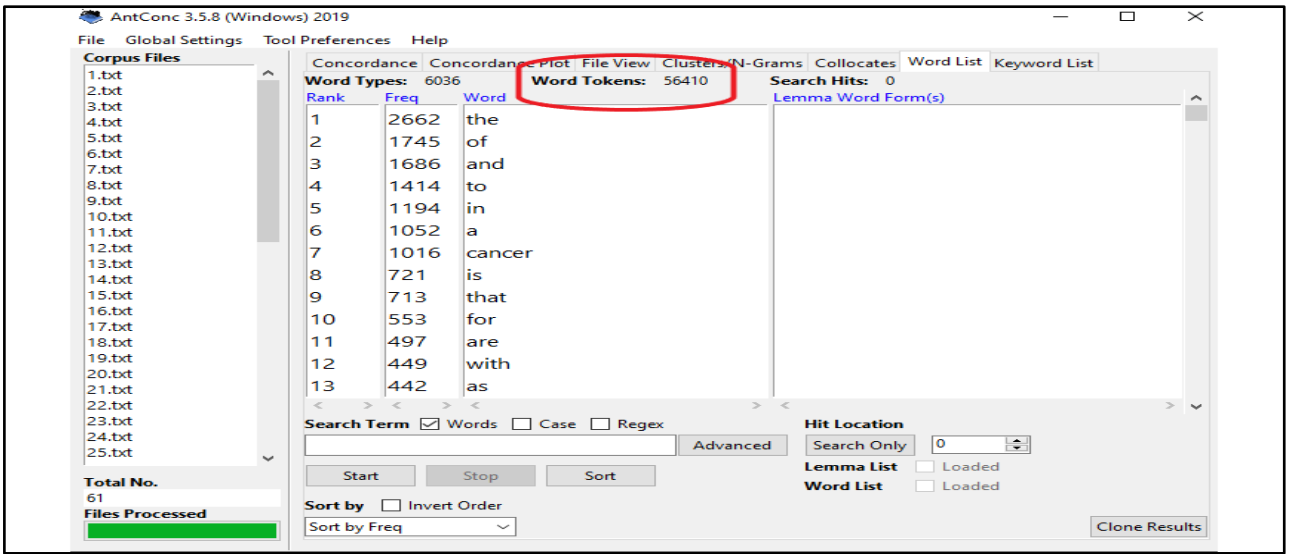

FIGURE 4. The size of the corpus

ANALYSIS AND RESULTS

SP comprises a conceptual representation of interaction between an *antagonist* (ODC) and a *protagonist* (IDC). The antagonistic entity, which embodies negative values, is construed as invading the protagonist's spatial ground. The antagonist's physical impact negatively affects the protagonist (IDC). Unless a preventive action is performed by the IDC, such impact is construed as immediate and destructive. The distinction between the center and the periphery elements encompasses geographical perspectives (Cap, 2013: 74). Since there are two kinds of entities in the DS (IDCs and ODCs), the first two categories accommodate the linguistic items that denote these entities. Category 1 of the SP strategy consists of NPs construed as deictic center elements of the DS (IDCs) that are represented by people and their health and safety. Category 2 consists of NPS construed as elements outside the deictic center of the DS as ODCs that are represented by carcinogens. The corpus analysis of these two categories has been performed according to the following steps:

1. Generating a wordlist by the Word List tool in AntConc as shown in Figure (5);

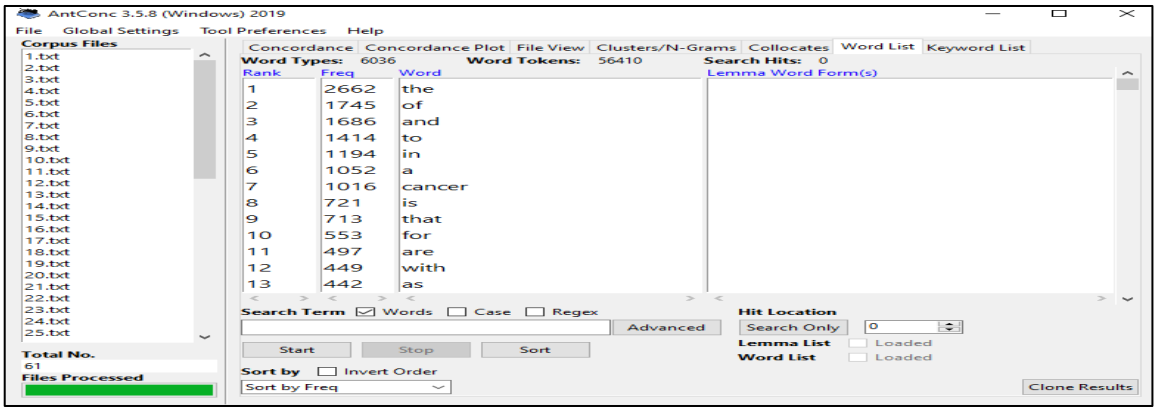

FIGURE 5. The wordlist of the Corpus

2. Investigating the nouns in the wordlist manually by the Concordance tool to identify the nouns (with their frequencies) that are related to people and their health and safety (as IDCs) and carcinogens (as ODCs) as shown in Figure (6):

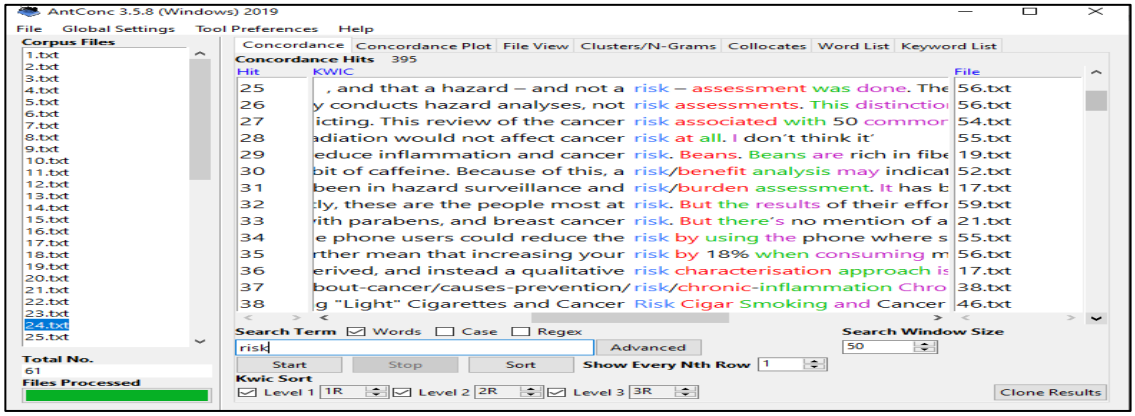

FIGURE 6. The Concordance investigation of the NPs in the word list

In few cases, concordances have not provided clear indication of the contexts of certain tokens. Therefore, the File View tool (which opens the file where a certain selected NP is located) is used, as shown in Figure (7):

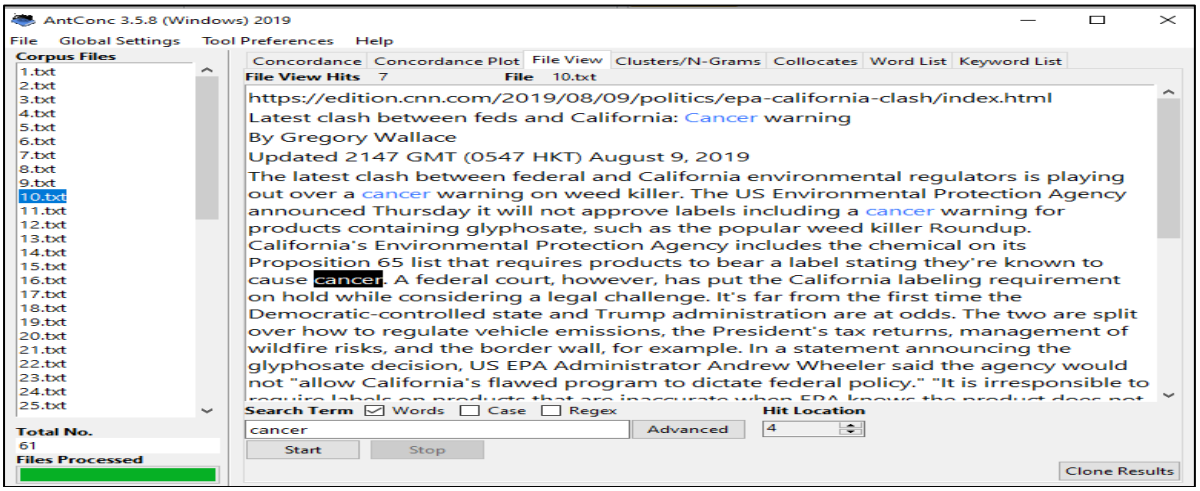

FIGURE 7. The File View tool

3. Using the Word List and the Cluster/ N-gram tools, tow lists of NPs have been generated: the first consists of NPs construed as IDCs and the second consists of NPs that are construed as ODCs. The NPs pointed out by Cluster/ N-gram tool have been identified using key adjectives (as search terms in NPs). These key adjectives are *carcinogenic, environmental, toxic, contaminated, lifestyle* and *electromagnetic* for the ODCs and *healthy* and *safe* for the IDCs. The reason behind using these adjectives in particular is that they have appeared in the word list of the corpus. For example, Figure (8) shows the Cluster/ N-gram investigation for the adjective *carcinogenic*:

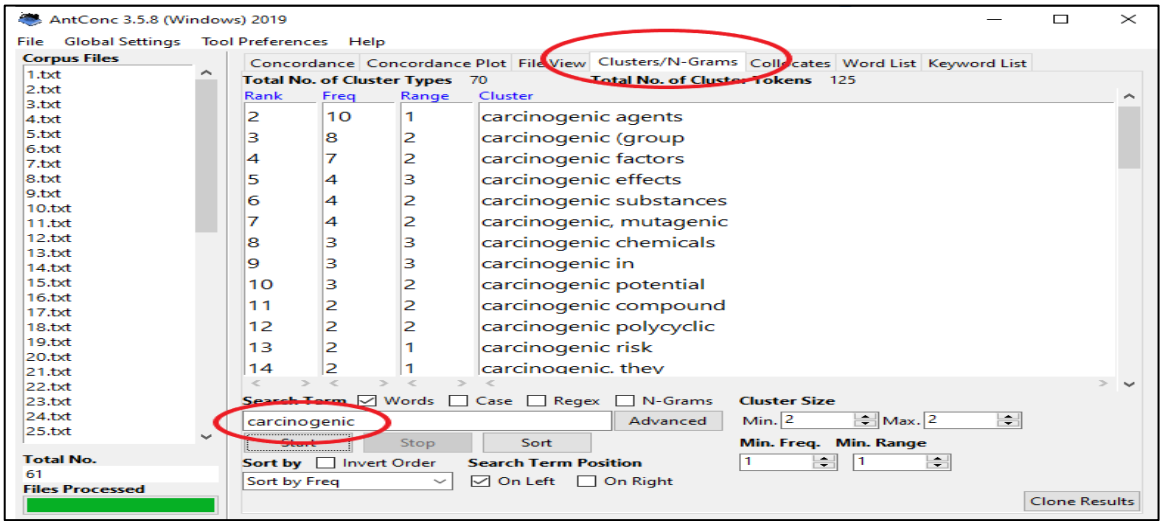

FIGURE 8. The Cluster/ N- gram investigation of certain adjectives

The Cluster/ N-gram lists have been manually filtered to exclude the irrelevant NPs. The final NP lists are shown in table (2):

TABLE 2. The NPs employed as IDCs and ODCs in categories 1 and 2

IDCs (category 1)	Frequency	ODCs (category 2)	Frequency
Health	151	Exposure	272
Safe*	74	Chemical(s)	182
People/ consumers	69	Carcinogen(s)	176
Woman/ women	33	Factor(s)	136
Girl(s)/ boy(s)	31	Meat(s)	135
Antioxidants	30	Food(s)	110
Man/ men	28	Smoke*	81
Worker(s)	21	Radiation	71
Safe+ behavior/ lifestyle etc.	13	Tobacco	69
College/ school+ students	10	Drug	69
they	10	Powder	58
		Alcohol	56
		Compound(s)	54
		Coffee	52
		carcinogenic	51
		Cigarette(s)	43
		Ranitidine	35
		NDMA	32
		Toxic chemical(s)/ compound(s)/ material(s) etc.	31
		Gas*	29
		Sunscreen(s)	29
		Sanitizer(s)	24
		Asbestos	23
		Contamination	23
		Dioxane	23
		Tattoo*	23
		HCAs	20
		HIV	20
		Charcoal	19
		it	19
		Glyphosate	19
		PAHs	19
		Benzene	18
		Toxin(s)	18
		Marijuana	17
		carcinogenic agents/ factors	16
		they	16
		Metformin	14
		Contaminated + bases/ equipment/ batches/ ground water/ soil etc.	14
		Obesity	11
		Hazard	10
		Pollution	10
		Electromagnetic field(s)/ spectrum	9
		Acrylamide	8
		Medication(s)	8
		Lifestyle choices/ habits/ factors/ changes	6

		Diabetes	5
		Zantac	5
		Bacon	5
Total instances of ODC NPs	**470**	**Total instances of ODC NPs**	**2193**

There are 470 instances for NPs that are construed as IDCs and 2193 NPs that are construed as ODCs.

Since it is impossible to precisely list all the IDC and ODC NPs in the table, some inferential means are used for denoting extra information in a synthetic manner to indicate a general lexico-grammatical context in which certain NPs are counted. Thus, the "*" wildcard represents one of the facilities that AntConc provides. It permits the investigation of different endings for the same stem at one round. Thus, *Smoke** in table (2) indicates that the count has included the NP *smoke* in addition to *smokes* and *smoking*. The "+" sing indicates that the adjective preceding this sign has occurred as a pre- modifier for a set of nouns which are all presented in the table and separated by slashes. For example, the adjective *contaminated* has occurred as a pre-modifier in NPs in which the heads are nouns like *bases/ equipment/ batches/ ground water/ soil* etc. The slash represents the choices available for the NP heads.

Category 3 consists of VPs "of motion and directionality construed as markers of movement of ODCs towards the deictic center" (Cap, 2013, p. 106). A motion verb encodes certain "types of semantic information: Manner of motion (e.g., hop), Cause (e.g., kick) and Path (e.g., exit, enter)" (Férez, 2008, p. 23). The element of Path can be encoded either by verbs or prepositions (e.g., *out, into*). Talmy (2000, p. 28; cited in Férez, 2008, pp. 30- 31) states that the "English expressions of Motion with conflated Manner or Cause" are as follows:

Move + Manner
Non-agentive
a. The rock *slid/rolled/bounced* down the hill
b. The gate *swung/creaked* shut on its rusty hinges
c. The smoke *swirled/squeezed* through the opening
Agentive
d. I *slid/rolled/bounced* the keg into the storeroom
e. I *twisted/popped* the cork out of the bottle
Self-agentive
f. I *ran/limped/jumped/stumbled/rushed/groped* my way down the stairs
g. She *wore* a green dress to the party
Move + Cause
Non-agentive
h. The napkin *blew* off the table
i. The bone *pulled* loose from its socket
j. The water *boiled* down to the midline of the pot
Agentive
k. I *pushed/threw/kicked* the keg into the storeroom
l. I *blew/flicked* the ant off my plate
m. I *chopped/sawed* the tree down to the ground at the base
n. I *knocked/pounded/ hammered* the nail into the board with a mallet

eISSN: 2550-2131
ISSN: 1675-8021

Directed motion verbs fall into two classes. The first class includes verbs associated with two-point scales in that they encode having or not having a particular feature. Such verbs are *arrive, depart, enter,* and *exit*. The second class includes multiple-point scales with many values. Such verbs are used for describing gradual traverse of the path. Such verbs are *advance, descend, fall, recede,* and *rise* (Rappaport Hovav & Levin, 2010, p. 30).

The corpus analysis of this part has been conducted by forming a list of the roots of all the regular motion and directionality verbs mentioned by Talmy (2000). Then each root has been separately investigated using the wildcard of * in Antconc to investigate different forms of the verb (base, with 3rd person singular –s, past, past participle and continuous), as shown in figure (9) for the verb *exist:*

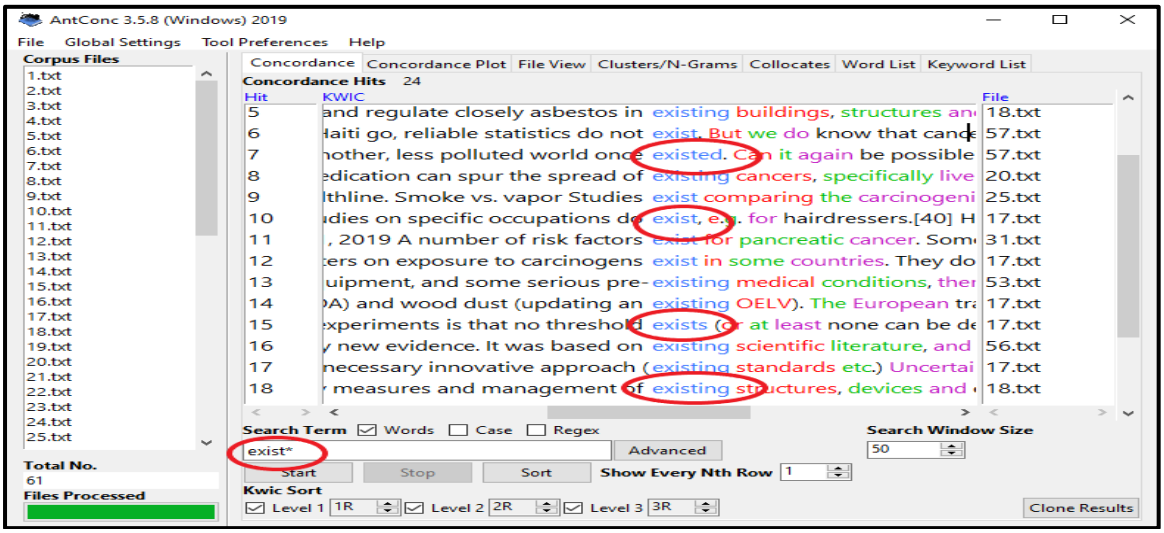

FIGURE 9. Investigating the verb "exist" using the wildcard *

As for irregular verbs, the root of each verb has been investigated separately with the wildcard * to include the base, 3rd person singular -s and continuous forms, as shown in figure (10) for the verb *wear:*

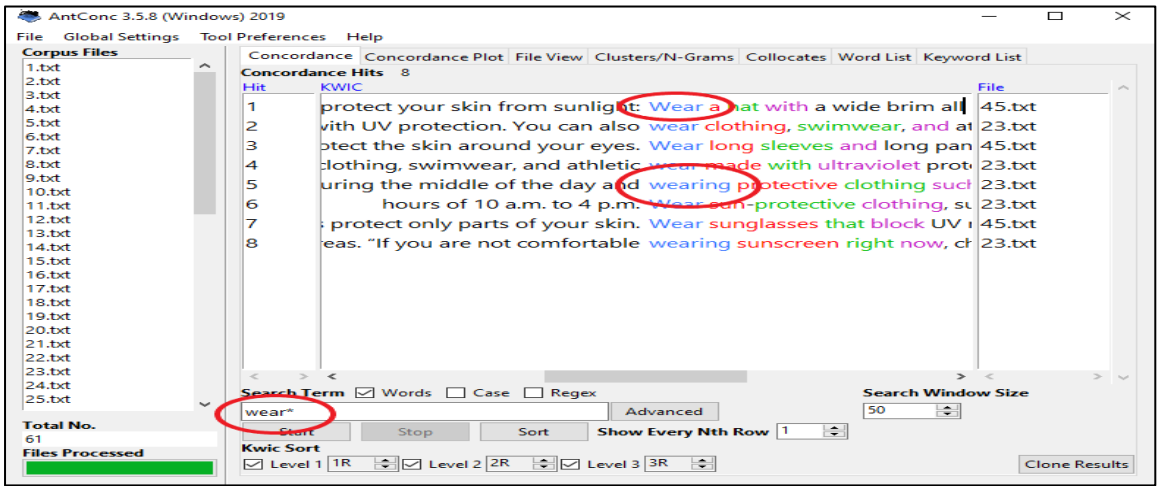

FIGURE 10. Investigating the verb "wear" using the wildcard *

Each form of the past and past participle for the irregular verbs has been separately investigated using the Concordance tool.

The concordance lists of all verbs have been filtered to exclude irrelevant cases. The File View tool has also been used when concordances provide limited view of the context. The final results are put in table (3):

TABLE 3. The distribution of motion and directionality VPs

VPs of motion & directionality	Frequency
Existed/ exist/ existing/ exists	8
Come+ in/ through/ with	8
Rise/ arise	7
fall	3
Go+ through/ up	3
raise	2
enter	1
kick	1
running	1
advanced	1
Total instances	**35**

According to Rappaport Hovav & Levin's (2010, p. 30) classification of motion and directionality VPs, the VPs in table (3) are of both categories of that classification: VPs associated with two-point scales in that they encode having or not having a particular feature and VPs used for describing gradual traverse of the path. Rearranging the VPs in table (3) according to this classification, table (4) results:

TABLE 4. Classification of motion and directionality VPs

VPs of tow- point scales (no gradual transvers of the path)	Frequency	VPs describing gradual traverse of the path	Frequency
Existed/ exist/ existing/ exists	8	Rise/ arise	7
Come+ in/ through/ with	8	fall	3
Go+ through/ up	3	raise	2
enter	1	kick	1
		running	1
		advanced	1
Total instances	**20 (57.2 %)**	**Total instances**	**15 (42.8 %)**

The VPs that do not elicit a gradual traverse have more frequency in the corpus (20 instances; 57.2 % for VPs with no gradual transvers and 15 instances; 42.8 % for VPs with gradual transvers).

Category 4 consists of VPs "of action construed as markers of impact of ODCs upon IDCs" (Cap, 2013, p. 107). Action verbs are the dynamic verbs that denote acts, events, or processes with an inherent sense of completion. Such verbs mark physical or communication abilities or actions that take place instantaneously (e.g., *shut, smash, swallow* and *throw*). An action verb may also present propositional information that is valid whenever the events or states actually occur (e.g., *lead* and *reveal*) (Biber, Johansson, Leech, Conrad & Finegan, 1999).

To analyze this category, Word List, Concordance and File View tools are used. The word list has been surveyed to point out the verbs that denote a dynamic sense and refer to an activity or a process. Then, the related verbs are hit to display their concordances in the corpus. Concordance helps figuring out whether the context indicates that the verb marks an impact of ODCs upon IDCs. The File View tool is used to search for more clues that provide more contextual elucidation. The results are shown in table (5):

TABLE 5. VPs of action

VPs of action	Frequency
Cause, causes, BE caused by	199
found	102
Use, used, be / HAVE been used, be+ using	101
Increase/ increases/ increased	98
increase	69
Exposed to+ NP	49
Smoke/ smoked	46
Be linked	21
(to) develop	21
(to) create/ creates/ created	19
Report, be/ HAVE (been) reported	17
Lead to	15
(modal auxiliary/ to) produce/ produces	15
Make/ makes	14
(To/ modal auxiliary) spread	14
Results, result + from/ in	13
(BE) Classified	13
Modal auxiliary/ to+ affect	13
Modal auxiliary be/ HAVE been/ BE + detected	12
HAVE been Recalled	10
BE/ HAVE been + formed	10
Get, BE getting	9
BE/ HAVE + smoked	6
BE cooking/ cooked	5
take	5
elevated	5
grill	4
Can be/ BE + affected	4
BE given	3
cook	2
Total instances	**914**

Category 5 represents the anticipated threat of the impact of the ODCs upon the IDCs and category 6 represents the catastrophe which comes as an effect of the impact of ODCs upon IDCs. The lexico- grammatical tools of categories 5 and 6 are NPs which denote abstract concepts. For

this reason, in addition to the subtle difference between them, the results of analyzing categories 5 and 6 are put together in table (6). Juxtaposing the results of analysis in one table helps readers more in figuring out the difference between these two categories and shows the significance of each category clearly. The corpus tools employed in the analysis of both categories are Word List, Concordances and, in some cases, File View. First, the NP which denotes anticipated threat or catastrophic effect is pointed out. Second, the concordance of such NP is checked to make sure that it is related to category 5 or 6. The File View tool has also been used.

TABLE 6. NPs of category 3 category 6

NPs of anticipated threat concepts (category 5)	**Frequency**	**NPs of catastrophe effect concepts (category 6)**	**Frequency**
Risk(s)	466	cancer	1143
Health/ increasing+ **concerns** + about…/ that…	17	(cancer) death/s	72
(cancer) **warning**	14	(Health/ toxic) effect/s	59
Health/ risk/ cancer+ **concern**(s)	12	disease	53
Strange/ warning+ **symptoms**	12	DNA/ lung/ cell/ etc. + **damage**(s)	22
Carcinogenicity	9	(Cancerous/ esophageal/ embryogenic/ malignant)+ **tumor**(s)	17
Genetic/ lethal/ spontaneous+ **mutations**	6	Cancer/ health/ contamination+ **issue**(s)	13
Danger	4	**Spread** of+ cancer/ cancer cells/ the tumors	9
carcinogenic potential	3	Leukemia	9
Warning calls	1	illness	5
		carcinogenic effects	4
		Serious consequences	1
Total instances	**544**	**Total instances**	**1407**

The ultimate statistical outcomes obtained from the application of the SP strategy are displayed in table (7):

TABLE 7. The Distribution of the SP strategy categories

No. of category	**Lexico- grammatical tools within the discourse space**	**Total instances**	**Percentage of instances**
1	Noun phrases (NPs) construed as elements of the deictic center of the DS (IDCs)	470	8.5%
2	Noun phrases (NPs) construed as elements outside the deictic center of the DS (ODCs)	2193	39.5%
3	Verb phrases (VPs) of motion and directionality construed as markers of movement of ODCs towards the deictic center	35	0.6%
4	Verb phrases (VPs) of action construed as markers of impact of ODCs upon IDCs	914	16.5%
5	Noun phrases (NPs) denoting abstract concepts construed as anticipations of impact of ODCs upon IDCs	544	9.7%
6	Noun phrases (NPs) denoting abstract concepts construed as effects of impact of ODCs upon IDCs	1407	25.2%
	Total instances in the English corpus	**5563**	**100%**

Table (7) shows that carcinogen risk has been spatially approximated in 5563 instances in the corpus. The table also shows the distribution of the six categories of SP strategy in the corpus. Figure (11) shows the graphic distribution of these categories:

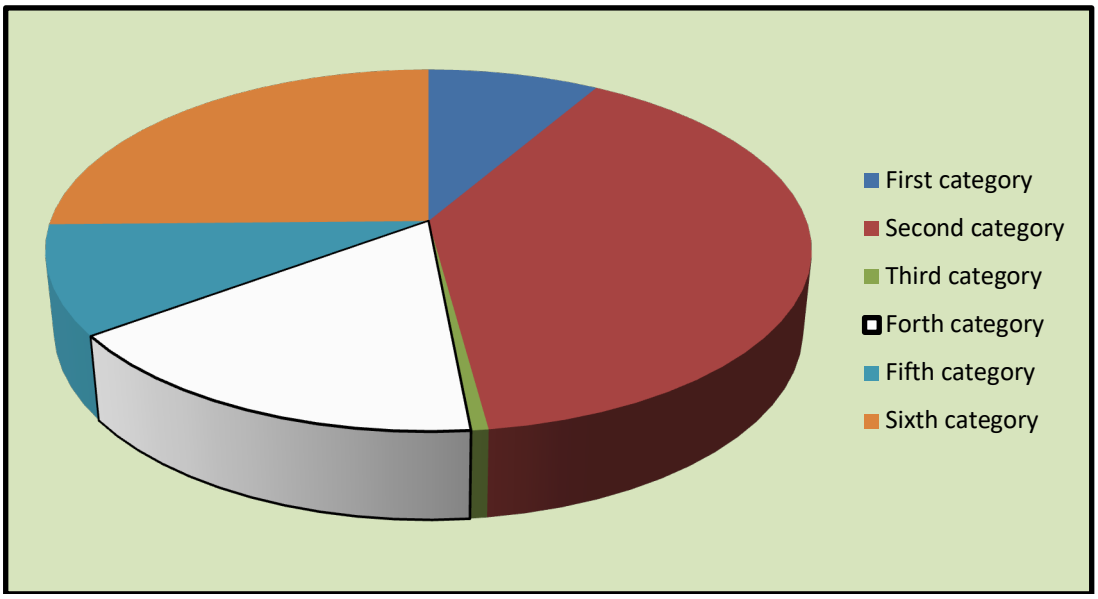

FIGURE 11. Distribution of SP strategy categories

DISCUSSION

All six categories of SP are used for constructing carcinogen risk. In this respect, Cap (2008, p. 32) asserts that the employment of all SP categories in discourse proves the pervasiveness of SP as a proximization tool. The corpus puts the categories of SP strategy to use in different rates. This finding is supported by Cap (2013, p. 108) who states that although SP exploits the largest amount of lexico-grammatical material in the proximization theory, this material is not all engaged at the same time. Table (8) displays the frequency and rank of the categories:

TABLE 8. The Ranks and frequency of the SP strategy categories

Category rank	Category NO.	Lexico- grammatical tools within the discourse space	Total instances	Percentage of instances
1st	2	Noun phrases (NPs) construed as elements outside the deictic center of the DS (ODCs)	2193	39.5%
2nd	6	Noun phrases (NPs) denoting abstract concepts construed as effects of impact of ODCs upon IDCs	1407	25.2%
3rd	4	Verb phrases (VPs) of action construed as markers of impact of ODCs upon IDCs	914	16.5%
4th	5	Noun phrases (NPs) denoting abstract concepts construed as anticipations of impact of ODCs upon IDCs	544	9.7%
5th	1	Noun phrases (NPs) construed as elements of the deictic center of the DS (IDCs)	470	8.5%
6th	3	Verb phrases (VPs) of motion and directionality construed as markers of movement of ODCs towards the deictic center	35	0.6%

eISSN: 2550-2131
ISSN: 1675-8021

Table (8) discloses the dominance of category 2 (2193 instances; 39.5%) with NPs construed as ODCs. Cap (2008, p. 32) assures that the high frequencies "determine the positions of the abstracted meanings mapped onto the [spatial] axis" in that "the higher the frequency counts …, the smaller… the distance to the deictic center". This fact suggests the closeness of carcinogens to human bodies and, therefore, reinforces the promotion of preventive measurements. The second rank goes to category 6 (1407 instances, 25.2%). According to this category, NPs denote abstract concepts which are construed as catastrophic effects of the impact of ODCs upon IDCs. What is special about the corpus under investigation is that the majority of NPs have construed physical rather than abstract effects of impact. This finding is supported by Cap (2006, p. 60) who assures that, in the context, the catastrophic effects can denote physical impact. Therefore, the effects of impact of carcinogens upon human bodies are mainly physical or physiological (concrete) (e.g. *leukemia, illness, damage, cancerous tumors*).

Pointing out the ODCs and their catastrophic effects, the corpus produces justifications for the catastrophic effects. This justification is fulfilled through category 4 which has occupied the third rank in the frequency scale (914 instances; 16.5%). VPs of action are construed as markers of impact of ODCs upon IDCs. The category that occupies the forth rank is category 5 in which certain NPs denote abstract concepts that are construed as anticipations of impact of ODCs upon IDCs. The category has scored 544 instances in the corpus (9.7%). The category that comes in the rank before the last (fifth rank) is the one in which the NPs are construed as IDC elements (category 1). This finding agrees with Cap (2008, 2013, and 2017). Obviously, there is a considerable difference between the frequencies of the NPs construed as ODCs within category 2 (2193 instances; 39.5%) and those construed as IDCs within category 1 (470 instances; 8.5%). While analyzing the corpus, it has been noticed that the ODCs take the semantic role of agents in sentences in which the patients are indications of the impact of the ODCs upon the IDCs, as shown in the following example from the corpus:

1. Benzene and sanitizers may cause cancer.

In addition, the corpus is loaded with passive sentences in which the subjects represent the anticipated risk (the eventual harmful effects or ODCs which are caused or formed by the NPs in the by- phrases) as is clear in (2):

2. Certain chemicals, called HCAs and PAHs, are formed when muscle meat, including beef, pork, fish, and poultry, is cooked using high-temperature methods.

In the last rank comes category 3 (35 instances; 0.6%) where VPs of motion and directionality are construed as signs of movement of ODCs towards the deictic center. The VPs that do not elicit a gradual traverse have more frequency in the corpus (20 instances; 57.2 % for VPs with no gradual traverse and 15 instances; 42.8 % for VPs with gradual traverse) than those that indicate gradual traverse.

CONCLUSION

The most salient conclusion is that spatial invasion is constructed with all of the six categories of the SP strategy. However, the corpus has employed the categories with different rates. The dominance of category 2 (NPs construed as ODCs) reveals the major role of scientific discourse in somaticizing as much carcinogen elements as possible. Such tendency comes within the essence of scientific discourse as an informative and knowledge constructing discourse. The linguistic status of SP is mainly the matter of ODC markers which act like the main initiators of the spatial shift from DS periphery to the DS center. This shift, in the SP architecture, forces the construal of risk which comes from an external physical impact on the central entities (IDCs) in the DS.

eISSN: 2550-2131
ISSN: 1675-8021

There is also considerable usage of NPs which are construed as catastrophic effects of the impact of ODCs upon IDCs (category 6). This fact goes in line with the dominance of the ODCs designation (category 2). The actors are after asserting the truth of the carcinogen effects of the ODCs. The ODCs designation is accompanied by equal designation of the catastrophic effects of the ODCs (upon the IDCs). The actors establish more convincing arguments that are supported by reasons why certain environmental entities are considered to be carcinogenic and fatal.

Lexico-grammatical tools from category 4 are used to justify the catastrophic effects created in category 6. ODC effects on IDCs are inferred from category 4 (VPs of action). As a result, the corpus illustrates how the catastrophic effect moves from the DS peripheral to its center. There is a shift away from arguing for the impact of carcinogens toward the establishment of the physical mechanism that carcinogens use to carry out their impact on human bodies.

The forth priority for the actors is to anticipate risk itself as a means for spatial invasion (category 5). However, this anticipation is not as prevalent in the corpus as the actions indicating the effects of the ODCs upon the IDCs or the indication of the catastrophic effects themselves (categories 4 and 6). Obviously, the actors prefer to focus on real cases and clinical instances rather than anticipating the impact of that invasion. Therefore, the focal point of the corpus is to warn about the ultimate catastrophic effects which have already been estimated in many people. The focal point that follows is the anticipation of the continuous spatial invasion of the risk that initiates from ever- lasting environmental carcinogens.

Spatial invasion does not show heavy reliance on IDCs since category 1 (NPs construed as IDCs) comes in the rank before the last. This reveals that risk heads from tremendous environmental carcinogens to human bodies in particular. Therefore, the number of IDC elements becomes limited in comparison with the number of the ODC elements. The corpus mainly concentrates on naming as much carcinogens and aspects of risk as possible at the expense of the IDCs which are well- known to interlocutors.

Spatial invasion is constructed in rapid and quick manner to launch sudden attacks. This tendency has boosted the legitimization of preventive actions of people to stay alarm of these sudden attacks. This conclusion is revealed by category 3 (VPs of motion and directionality construed as signs of movement of ODCs towards the deictic center) that scored the least frequency in the SP strategy. It also indicates that spatial invasion depends on the stable, potential and static carcinogens that surround people in the environment. Therefore, the speed of carcinogens is of no potential value to the conceptualization of spatial invasion.

The findings arrived at are of significance to scientific discourse producers whose awareness of the cognitive architecture of such discourse can be enhanced by the deep and detailed analysis and results. The corpus linguistic results can draw their attention to the statistical significance of certain linguistic tool in shaping scientific knowledge. Moreover, experts, researchers, and professionals of English are urged to look further into scientific discourse of medicine issues. Such studies can contribute to the understudied field of medical linguistics. In highlighting the uniqueness of scientific discourse in general and medical discourse in particular, their findings are anticipated to be useful to medical and scientific specialists, experts, and columnists.

In the light of the findings obtained, further studies can be conducted to investigate the construction of other critical environmental physical issues and threats in scientific discourse by applying Cap's (2013) proximization theory such as global warming, drought, pandemics, famine, etc.

eISSN: 2550-2131
ISSN: 1675-8021

REFERENCES

Aajami, R. F. (2020). Cognitive Implications of Usage-Based Approach. *Journal of the College of Education for Women*, *31* (44), 1-13.

Abbas, N. F. (2020). Pragmatics of Overlapping Talk in Therapy Sessions. *Journal of Language and Linguistic Studies*, *16* (3), 1251-1263. Doi: 10.17263/jlls.803705

Adegbite, W. (2000). "Pragmatics: Some Basic Principles". In A.O. Babajide (Ed.), *Studies in English Language*. Ibadan: Enicrownfit Publishers, pp. 60-76.

Anthony, L. (2019). AntConc. Retrieved January 7, 2022, from http://www.laurenceanthony.net/software/antconc/.

_____ (2019). AntFileCnverter. Retrieved January 7, 2022, from http://www.laurenceanthony.net/software/antconc/.

Biber, D.; Johansson, S.; Leech, G.; Conrad, S. & Finegan, E. (1999). *Longman Grammar of Spoken and Written English*. London: Longman.

Brezina, V. (2018). *Statistics in Corpus Linguistics: A Practical Guide*. Cambridge: Cambridge University Press.

Cap, P. (2005). Language and Legitimization: Developments in the Proximization Model of Political Discourse Analysis. *Lodz Papers in Pragmatics, 1*, 7–36.

_____ (2006). *Legitimization in Political Discourse: A Cross-disciplinary Perspective on the Modern US War Rhetoric*. Newcastle: Cambridge Scholars Press.

_____ (2008). Towards the Proximization Model of the Analysis of Legitimization in Political Discourse. *Journal of Pragmatics, 40*, 17–41

_____ (2013). *Proximization: The Pragmatics of Symbolic Distance Crossing*. Amsterdam: John Benjamins.

_____ (2014). Applying Cognitive Pragmatics to Critical Discourse Studies: A Proximization Analysis of Three Public Space Discourses. *Journal of Pragmatics*, *70*, 16- 30.

_____ (2017). *The Language of Fear: Communicating Threat in Public Discourse*. London: Palgrave Macmillan.

_____ (2018). Spatial Cognition. In J. Flowerdew and J. Richardson (Eds.) *The Routledge Handbook of Critical Discourse Studies* (pp. 92–105). London: Routledge.

_____ (2020). Representation, Conceptualization and Positioning in Critical Discourse Analysis. *International Review of Pragmatics*, *12*, 272–294.

Carcinogen (2008). *Webster's New World Medical Dictionary.* New Jersey: Wiley Publishing, Inc.

Chilton, P. (2005). Discourse Space Theory: Geometry, Brain and Shifting Viewpoints. *Annual Review of Cognitive Linguistics*, *3*, 78–116.

Denholm, R.; Schüz, J.; Straif, K.; Ali, F.M.H.; Bonas, F.; Gjebrea, O.; Sifton C. & Olsson, A.C. (2016). Environmental carcinogen exposure and lifestyle factors affecting cancer risk in Qatar: findings from a qualitative review. *Eastern Mediterranean Health Journal*, *22* (3), 219- 227.

Diniz, L. (2005). Comparative Review: Textstat 2.5, Antconc 3.0, and Compleat Lexical Tutor 4.0. *Language Learning & Technology*, *9* (3), 22-27.

Dunmire, P. (2011). *Projecting the Future through Political Discourse: The Case of the Bush Doctrine.* Amsterdam: John Benjamins.

Férez, P. C. (2008). *Motion in English and Spanish: A Perspective from Cognitive Linguistics, Typology and Psycholinguistics* (Doctoral thesis, Departamento de Filología Inglesa

Facultad de Letras Universidad de Murcia). Retrieved November 5, 2021, from https://www.tdx.cat/bitstream/handle/10803/10816/CifuentesFerez.pdf

Filardo Llamas, L. (2010). Discourse Worlds in Northern Ireland: The Legitimisation of the 1998 Agreement, in K. Hayward and C. O'Donnell (Eds.), *Political Discourse and Conflict Resolution Debating Peace in Northern Ireland*. London: Routledge, pp. 62 – 76.

Gatto, N. M. (2021). Environmental Carcinogens and Cancer Risk. *Cancer*, *13*, 1- 2. doi.org/10.3390/cancers13040622. Retrieved November 11, 2021, from https://www.mdpi.com/journal/cancers

Goh, C. F., Chan, M. Y., Ali, A. M. & Rashid, S. M. (2019). The Complementary and Alternative Medicine (CAM) Product Information Brochure: How is Generic Structure Used to Persuade Potential Users? *GEMA Online® Journal of Language Studies*. *19* (4), 219- 242. http://doi.org/10.17576/gema-2019-1904-12

Gustilo, L, May Pura, C. & Biermeier, T. (2021). Coronalexicon: Meanings and Word-formation Processes of Pandemic-related Lexemes across English Varieties. *3L: Language, Linguistics, Literature® The Southeast Asian Journal of English Language Studies*, *27*(4), 1- 15. http://doi.org/10.17576/3L-2021-2704-01

Hanauer, D. I. (2006). *Scientific Discourse Multiliteracy in the Classroom.* London: Continuum.

Hart, C. (2010). *Critical Discourse Analysis and Cognitive Science: New Perspectives on Immigration Discourse*. Basingstoke: Palgrave Macmillan.

Idris, H. & Ghani, R. A. (2012). Construction of Knowledge on Facebook. *3L: Language, Linguistics, Literature® The Southeast Asian Journal of English Language Studies*, *18*(3), 61 – 72.

Jaafar, E. A & Jasim, H. A. (2022). A Corpus- Based Stylistic Analysis of Online Suicide Notes Retrieved from Reddit, *Cogent Arts & Humanities*, *9*(1), 1-16. DOI:10.1080/23311983.2022.2047434

Johnstone, B. (2008). *Discourse Analysis*. Singapore: Blackwell Publishers.

Kaal, B. (2012) Worldviews: Spatial Ground for Political Reasoning in Dutch Election Manifestos. *CADAAD*, *6* (1), 1 – 22.

Khalil, H. H. & Al- Zubaidi, N. A. G. (2022a). Constructing Carcinogen Risk in Scientific Discourse through Ideological Conflict: A Cognitive Pragmatic Analysis. *Theory and Practice in Language Studies, 12*(8), 1489-1499. DOI: https://doi.org/10.17507/tpls.1208.04

Khalil, H. H. & Al- Zubaidi, N. A. G. (2022b). Constructing Imminent Carcinogenic Attack in English and Arabic Scientific Discourse: A Corpus-Based Contrastive Study. *Journal of the College of Education for Women, 33*(4), 1- 18. DOI: https://doi.org/10.36231/coedw.v33i4.1631

Lewis, M. (2001). *Teaching Collocation: Further Developments in the Lexical Approach*. London: Commercial Colour Press.

Mordovina, T. V. & Nikulshina, N.L. (2010). From Analysis to Teaching Types of Foreign Language Scientific Discourse. Retrieved October 7, 2021 form https://www.tstu.ru/en/science/st/pdf/2010/mordovina.pdf

Myers, G. (2003). Discourse Studies of Scientific Popularization: Questioning the Boundaries. Discourse Studies, *5*(2), 265–279. https://doi.org/10.1177/1461445603005002006

Okulska, U. & Cap, P. (Eds.) (2010). *Perspectives in Politics and Discourse*. Amsterdam: John Benjamins.

Overton, J. A. (2013). "Explain" in scientific discourse. *Synthese*, *190*(8), 1383–1405. http://www.jstor.org/stable/41931908

Paltridge, B. (2000). *Making Sense of Discourse Analysis.* Brisbane: Antipodean Educational Enterprises.

Rappaport Hovav, M., & Levin, B. (2010). Reflections on manner/result complementarity. In M. Rappaport Hovav, E. Doron, & I. Sichel (Eds.) *Lexical Semantics, Syntax, and Event Structure* (pp. 21–38). Oxford: Oxford University Press.

Selvaraj, S. & Sandaran, S.C. (2019). Discourses of Flood Disaster Preparedness by NGOs: Humanitarian Aid, Teamwork and Victimization. *GEMA Online® Journal of Language Studies*. *19* (4), 111- 127. http://doi.org/10.17576/gema-2019-1904-06

Skier, J. & Vibulphol, J. (2016). Development and Use of a Corpus Tailored for Legal English Learning. *PASAA*, *25*, 237- 254. Retrieved November 2, 2021, from https://files.eric.ed.gov/fulltext/EJ1134680.pdf

Tabbert, U. (2016). *Language and Crime: Constructing Offenders and Victims in Newspaper Reports.* London: Palgrave Macmillan.

Talmy, L. (2000). *Toward a Cognitive Semantics (Vol. 2): Typology and Process in Concept Structuring*. Cambridge: The MIT Press.

Van Dijk, T. A. (2002). Political discourse and political cognition. In P. A. Chilton, & C. Schäffner (Eds.), *Politics as Text and Talk. Analytical Approaches to Political Discourse* (pp. 203-237). Amsterdam: John Benjamins. https://doi.org/10.1075/dapsac.4.11dij.

Wei, M. & Yu, G. (2019). On the Characteristics of Scientific Discourse and Translation. *Theory and Practice in Language Studies, 9* (8), 946-950. DOI: http://dx.doi.org/10.17507/tpls.0908.08

Younus, L. L. (2020). Audience Demand in Father-Son Argumentation across Generations. *Al-Adab Journal*, *134*, 13- 40.

Yu, G. (2009). Cohesive Coherence in Text and Translation. *Language and Translation*, *4*, 47-51.

ABOUT THE AUTHORS

Huda H. Khalil is an Assoc. Prof. in the University of Baghdad. She has been teaching English language and linguistics related subjects in the University of Baghdad for B.A and M.A students since 2002.

Nassier A. G. Al-Zubaidi is an Assoc. Prof. of English language and linguistics at the Department of English of the Faculty of Arts in Baghdad University, Iraq. He has a Ph.D. degree in English language and linguistics.