Compiling a cognition-based thematic monolingual lexicon

The case of violence

Esra' Moustafa Abdelzaher
Ain Shams University (Egypt)

This study falls within the scope of cognitive lexicography which uses cognitive linguistic theories in lexicographic practice. The main objective of the study is to create a cognition-based monolingual thematic lexicon. The lexicon tests the validity of using cognitive linguistics, which uses language to reveal the human perception of a concept, in defining controversial multidisciplinary concepts. To that end, violence is selected as a case study and FrameNet is recruited as a cognitive linguistic resource. Cambridge Smart Thesaurus and WordNet are used as secondary resources to FrameNet. English TenTen corpus is employed to authenticate the findings before placing them in the lexicon. A twelve-frame lexicon is the result of the study. The constructed lexicon linguistically includes more than 250 violence-expressing word senses, defined and placed within their violence-associated frames. Some frames are cited from FrameNet without modification, while others are conceptually and linguistically modified. More important, some violence-specific frames are newly-reported. Evidently, studying how physical violence is linguistically expressed displays how the concept is structured in the human cognition. Thus, an empirical cognition-based definition of violence is suggested. This meets the challenge of the multiple sociological, psychological, political and criminological definitions. Moreover, a comprehensive definition of violence is recommended to include both its associated frames and expressing words.

Keywords: cognitive lexicography, corpus linguistics, language of violence, FrameNet, WordNet

1. Introduction

Adopting a cognitive linguistic approach in the process of constructing or developing dictionaries is referred to as Cognitive Lexicography (CL). Ostermann

(2015), expounding CL, argues that using the theories of cognitive linguistics in lexicographic practice enhances the capacity and comprehension of dictionary information. In a case study of emotion terms in *Oxford Advanced Learner's Dictionary*, Ostermann (2012) highly recommends cognition-based modifications of their definitions. Thus, she uses FrameNet (FN)'s 'Feeling' frame and Kövecses' prototypical emotion scenario to form the new definitions.

Comparing the OALD's definition of happy to the cognition-based new definition of happiness reveals the effectiveness of the cognitive linguistic approach to lexicography. Happy is defined by OALD as a 'feeling or showing pleasure'. The 'Feeling' frame and the prototypical emotion scenario render a more detailed definition. Happiness is cognitively defined as 'the good and vital feeling when everything in your life is the way you want it and you don't have any problems; it might make you feel as if you were in heaven, you smile and everything is ok' (Ostermann 2015).

This study starts from the fact that language can be used to 'reflect patterns of thoughts' and explore a concept's structure in the human mind (Evans 2006). It is motivated by the fruitfulness of the cognitive linguistic approach to lexicography. It aims at building a cognition-based lexicon addressing one multidisciplinary notion: violence. The study recruits FN data and methodology to present a lexicon of violence. The proposed lexicon is sought to gather violence-expressing words in English and violence-associated frame(s) in the human mind. It, thus, tests the ability of CL to meet the challenge, of defining violence, created by the various sociological, psychological, political and criminological definitions. FN is extensively employed for its leading successful application of a cognitive linguistic theory: Frame Semantics, in building a lexical resource: FrameNet itself.

2.   Previous studies

FN has been widely and multilingually used to build thematic dictionaries. Schmidt (2009) presents one of the first and most famous attempts of constructing a FN-based dictionary covering language of football. The proposed dictionary: Kicktionary, covers English, French and German languages. Similar to FN, Kicktionary starts with a typical football scene, sets some potential participants, and suggest a list of typical evocative words. Then, it explores the suggested frame in a corpus to create the final frame, based on authentic data. Kicktionary follows the FN methodology, except for using an English general reference corpus. It enables domain-specific multilingual corpora compiled from English, French, and German football reports. It also uses WordNet (WN), which is a hierarchical lexical resource, to group the frame evocative words in synonymy sets. Therefore,

Kicktionary provides its users with cognitive information, i.e., frames, and ontological classification, i.e., hyponyms, of a target word. This enables easier translation, summarization and paraphrase from and into the three target languages. Kicktionary, according to Schmidt (2007), provides its users with 2000 multilingual LUs consolidated by more than 8000 authentic sentences.

Bertoldi and Chishman (2011) also use FN to build a domain-specific language resource addressing the legal language in the Brazilian system. However, they depend on FN data as a starting point, hypothesizing that frame structures are cognitively universal. Their lexicon uses the English Criminal_process frame as a nucleus and tries to expand it to the Brazilian Portuguese language. However, the different pragmatics of the Brazilian legal system from the American one impedes the full expansion process. Therefore, the expansion methodology is applied as follows: (a) maintaining frames in case of perfect frame match; (b) partial use in case of untypicality; (c) exclusion in case of absence in the target language or (d) modification to match the target language.

Moreover, a language, attempting to create a FN project, may start with a certain theme or domain, in the early stages of construction. Spanish FN initially targets the themes of communication and emotion in a 300-million-word corpus of Spanish texts. It employs FN methodology, annotation tools and database structure. Subirats and Petruck (2003) report the creation of a set of frames covering the emotion theme. This domain-specific work enables cross-cultural and cross-linguistic exploration of concepts. Thus, more cognitive commonalities are revealed and the role of pragmatics in cognition is identified.

Similar to the above-reviewed studies, this study constructs an FN-based domain-specific lexicon of violence. It integrates, as Schmidt's (2009) Kicktionary, other lexical resources with FN. In the present case, they are Cambridge Smart Thesaurus and WN. However, it does not start from the scratch. The presented lexicon, like the Brazilian Portuguese one (2011), depends on FN data, yet there is no target language. Only the English language is targeted in this attempt. Thus, instead of comparing the constructed frames cross-linguistically as Subirats and Petruck (2003), this study compares the suggested new frames to FN's original violence frame(s).

3. Theoretical framework

3.1 FrameNet structure

FrameNet is theoretically-based on Fillmore's Frame Semantics (FS) theory (1975). FS hypothesizes that each word sense is cognitively associated with a certain

'frame' stored in the mind of language users. This frame represents the typical situation in which this word sense is employed. Whenever the word is used, the situation is evoked and whenever the situation is present, the word is summoned too. Fillmore, Johnson and Petruck (2003), stressing the conceptual significance of a frame to language, describe a frame as the abstract keystone, of a lexicon, that facilitates grouping variable senses of a word together. The FS approach is described, by Atkins and Rundell (2008), as the most 'helpful' and 'appropriate' way to study a corpus.

FN is a machine-readable frame-based lexicon. It covers more than 13000 word senses and 1200 frames, according to the official FN website. FN names a word sense a 'lexical unit', henceforth LU, (Baker, Fillmore and Cronin 2003). De Cao, Croce and Basili (2010) point out that a LU can be a noun, verb, adjective or a multiple-word construction. To elaborate, 'assault' as a verb, 'assault' as a noun, 'abusive', and 'domestic violence' are LUs in the Abusing frame. Typically, a frame is evoked by several LUs and a word is capable of evoking multiple frames based on its different LUs. Atkins, Rundell and Sato (2003) clarify that different LUs instantiated by the same orthographic word occur within different syntactic structures and with different lexical collocates. This is the way various word senses are identified in a corpus. A frame can be thought of as a container of cognitively-related LUs.

As illustrated, FN places each LU within the frame it evokes. It also provides a full definition of the frame and a clear identification of its participants. Frame participants are labeled Frame Elements (FEs) in FN. FEs are somehow equivalent to the traditional semantic roles: agent, patient, experiencer, goal, etc., which are played by the verb arguments. However, semantic roles are more situation-based in FrameNet, (Fillmore, Petruck, Ruppenhofer and Wright 2003). FEs, as illustrated by Fillmore and Baker (2010), are either core or peripheral. Core FEs are inevitably present when the frame is activated and they contribute to the understanding of the frame. Peripheral FEs, however, may be instantiated to communicate time, place, frequency, and manner information. Moreover, Ruppenhofer et al. (2016) clarify that an FE may require or exclude the presence of another FE. For instance, in the 'Similarity' frame, 'entity1' requires the existence of 'entity2', while the existence of 'entity1' and 'entity2' excludes the presence of 'entities' as a core FE. That is to say, the participants in the 'Similarity' frame are either 'entity1' and 'entity2' or 'entities'.

3.2 FrameNet cognitive integrity

One of the innovative FN contributions to cognitive lexicography is its Frame-To-Frame (FTF) relations. Beside classifying words cognitively into frames, FN investigates such frames to reveal more cognitive associations among them. FN names

seven types of FTF relations: inheritance, subframe, uses, see also, perspective, causes, and inchoative of. Only two of them are discussed in this study because they are used in the lexicon of violence, for more details see (Ruppenhofer et al. 2016).

First, frames can be related to each other by the 'inheritance' relation. Baker et al. (2003) define 'inheritance' as a relation between a 'parent' frame and 'child' frame. The child frame inherits the semantic and conceptual characteristics from its parent. Each FE in the child frame corresponds to a parent FE. However, the parent frame may have core FEs which are changed into peripheral FEs in the child frame. The 'Killing' frame, as a demonstration, has 'killer', 'instrument', 'means' and 'victim' as core FEs. It is inherited by 'Execution' which has only two core FEs: 'executed' and 'executioner', while 'instrument' and 'means' are peripherally present. The significance of this relation is attributed to its correspondence to the *is-kind-of* ontological relation. It plays the most effective role when integrating FN database to WordNet and building hierarchical dictionaries.

Second, a parent frame cane be related to a child frame by a 'use' relation. Ruppenhofer et al. (2016) state that sometimes a frame presupposes the knowledge of a more general abstract frame. The child frame 'uses' a parent frame to be comprehended. For example, 'Violence' and 'Cause_harm' are related to each other through the 'uses' relation. 'Violence' presupposes the knowledge of the more general and abstract frame 'Cause_harm'. That is to say, full understanding of 'Cause_harm' is essential to grasp the concept of violence.

The lexicon of violence is structured after FN. It treats each word as a set of LUs and groups LUs in their respective frames. It defines each frame and clearly states its FEs in terms of coreness and periphery. It links the proposed frames to each other through inheritance and use relations. Moreover, the lexicon provides annotated examples illustrating the linguistic realization of each frame, especially the newly-suggested ones.

4.   Methodology

FN adopts a practical corpus-based methodology to create new frames. It starts with targeting a concept, then setting a working definition for it, and listing as many representative words as possible. Finally, it checks the use of these words in a corpus and annotates sample sentences in order to write the final definition of the frame and determine, by adding or excluding, its evocative words, (Baker et al. 1998; Fillmore & Baker 2010).

Building the Lexicon of Violence, in this study, goes through two stages. In the pre-lexicon stage, FN methodology is exploited with a slight change. First, the concept of violence is targeted, but instead of composing a working definition,

FN's definition of violence is operationally recruited. Then, to secure a comprehensive understanding of violence, its cognitively related frames in FN are traced.

Next, a list of all evocative words is formed and processed via Cambridge Smart Thesaurus which provides a number of corpus-based topics for each word. Searching a LU in CST returns multiple topics covering the target LU and other LUs instantiated by the same word. So, only violence-related topics are selected. Categorizing the LUs under CST topics and comparing the new classification to that of FN results in the creation of new violence-related frames. Based on FN violence frame and CST suggested topics, a cognition-based definition of violence is formed. This step initiates the process of writing the lexicon entries as it identifies the core FEs of the violent scene and facilitates linking different LUs and newly-constructed frames to each other.

After creating a frame, defining its FEs and determining its LUs, WN is used to gather the synonyms, if any, of each LU. WN is used for its unique thesaurus and ontological features. It provides a super-ordinate term covering a set of synonyms and presents different synonyms for each word sense. Moreover, its word sense differentiation is fine-grained, similar to that of FN. Definitions of a word sense in WN and FN, when not identical, are very close to each other. For each LU, FN definition is compared to the different definitions provided by WN. Then, the synonyms of the closet, sometimes typical, definition are added to the lexicon. WN is proven to facilitate accurate gathering of LU synonyms. This step allows hierarchical grouping of LUs in the same frame. The methodology of constructing the lexicon can be visualized as follows:
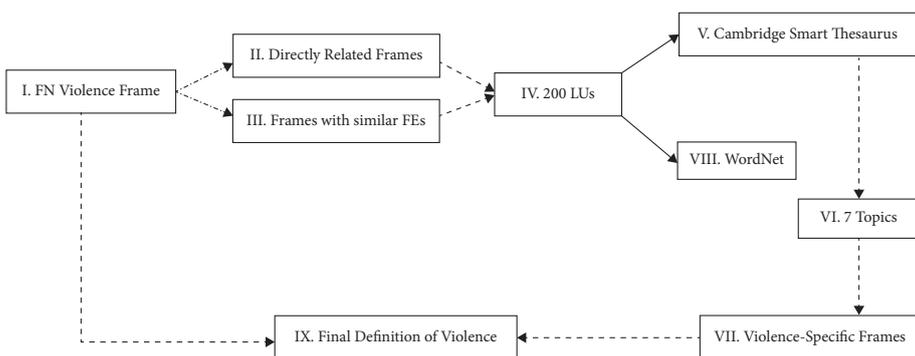

Figure 1. Lexicon of Violence Construction Methodology

5.   Analysis of data

5.1   Analyzing violence-related data in FN

In the preliminary stage, FN data is mined to collect any data relevant to the concept of violence. Searching violence in FN returns a single result pairing a LU and a frame. To elaborate, violence is a single-sense word evoking a frame carrying the same name. Violence as a LU is defined as 'an act that causes harm or injury upon another'. It typically evokes a frame of 'acts (or situations characterized by acts) that cause injury or harm. The acts may involve an Aggressor or Cause injuring a Victim, or Aggressors causing each other harm'.

FN's definition of violence, when deeply investigated, reveals valuable cognitive richness. The definition is supposed to reflect how violence is perceived and conceptualized in the human mind. First, it refutes the claimed sociological neutrality of violence by its negative labelling of the participants and the result of the act. Second, it implicitly refers to two types of violence: one-sided and mutual. It clearly states that a violent act may be committed by an aggressor against a victim or by aggressors against each other. Third, it attempts to cover both intentional aggressive violence of people and unintentional violence of natural powers. Four metalinguistic situations can be identified in this definition: (a) natural forces destruction; (b) human violence; (c) attack and (d) fight. The first scene is out of the present study scope. The other three physical violence scenes are covered in the proposed lexicon. Aggressor, aggressors, cause and victim are recorded as core FEs in the frame of 'Violence'.

Despite the richness of the definition, only one frame is provided as a cognitively-associated: 'Cause_harm'. 'Violence' uses 'Cause_harm' as a more general frame essential to understanding what violence is. 'Cause-harm', limits the scope of the violent act and the inflicted injury. The provided annotated examples in this frame prove that the kind of harm is physical and is directed to body parts, which is a core FE in this frame, along with agent, cause, and victim. Moreover, it 'uses' 'cause_bodily_harm' frame which crystallizes that the kind of harm intended herein is physical.

'Cause_harm', as a more abstract parent frame to Violence, is investigated to identify its related frames and determine how (dis)similar they are to Violence. It is proven to be linguistically and cognitively rich. Its network of related frames includes; 'Abuse', 'Terrorism', 'Toxic_substance' and 'Experience_bodily_harm'. Then, these frames are searched for in FN to identify their related frames, if any. Figure 2 illustrates the resulted violence-related frames of this bootstrapping process.
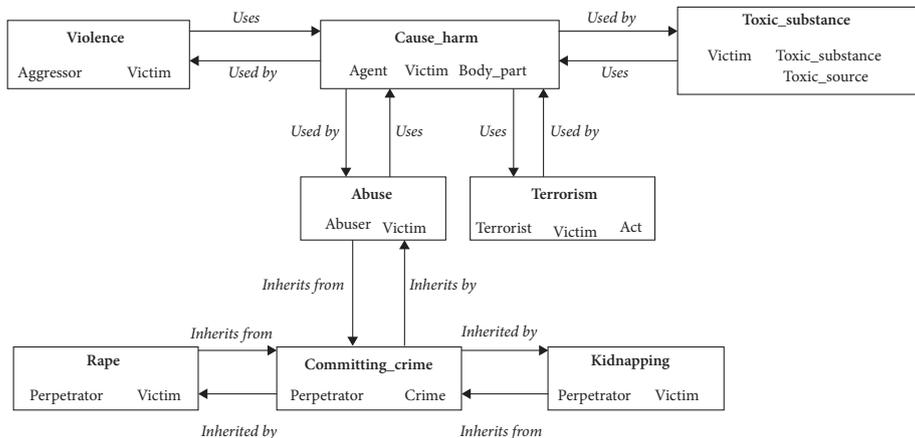
Figure 2. Violence_Related Frames in FrameNet

5.2 Excluding and including frames

The seed and bootstrapped frames are explored in order to conclude a similarity pattern, if any. As illustrated in Figure 2, there is an implied similarity between the FEs of the seed frames. All frames typically refer to an act committed by an agent against a victim. This agent is sometimes called aggressor or perpetrator and the victim can be labeled experiencer. Moreover, the definitions of the frames refer to some kind of harm, whether explicit or implicit. For instance, 'Abuse' is defined as 'an Abuser repeatedly treats the Victim in a cruel and violent way, including physically harming…'. Not only physical harm is mentioned but also violence is explicitly indicated as a manner of conducting the act.

Accordingly, two similarity patterns are identified to include frames as being violence-related or exclude them. First, if the core FEs include aggressor or aggressors and victim, or analogous FEs, and there is an implied or explicit physical harm involved in the definition of a frame, it is included. Second, if the previously-mentioned FEs are present, yet harm is totally absent, explicit reference to violence in the frame definition is essential to include the frame. So, the whole FN data is explored to capture frames having these patterns.

Being related to a violence-related frame does not guarantee inclusion in the lexicon. 'Legality', for instance, is bootstrapped as a frame using 'Commiting_crime'. It, however, does not refer to any of the target elements whether explicitly mentioned in its identified core elements or implicitly referred to in its definition. 'Robbery' is also excluded, despite being inherited by 'Commiting_crime' and having perpetrator and victim as core elements, because the implied damage/harm, in this case, is not physical. It is related to loss of property not injury of a

body part. Every criminal situation is cognitively associated with perpetrator and victim but it is not always linked to physical violence.

Hence, the recruited frames are selected according to three major criteria; (i) being directly related to the frame of violence by one of the FN set of relations; (b) sharing parallel cognitive structure: FEs and harm or (iii) sharing parallel cognitive structure and stating violence in the definition of the frame. The cognitive structure meant herein is the aggressor/victim or the aggressors FEs. Table 1 displays the recruited frames and the inclusion criterion or criteria they meet.

Table 1. Recruited violence-related frames and inclusion criteria

Frame	Criterion	Criterion Realization
Cause_harm	Directly-related	Use relation
Experience_bodily_harm	Cognitive structure Harm reference	Experiencer and Body_part 'An Experiencer is involved in a bodily injury…'
Abusing	Cognitive structure Violence reference	Abuser and victim 'in a cruel and violent way, including physically harming…'
Terrorism	Cognitive structure Violence and harm references	Terrorist, victim and act 'commits a violent or otherwise harmful act…'
Rape	Cognitive structure Implicit harm	Perpetrator and victim
Killing	Cognitive structure Implicit harm	Killer and victim 'causes the death'
Attack	Cognitive structure Harm reference	Assailant and victim '…causing or intending to cause the Victim physical damage…'
Counter_attack	Cognitive structure Harm reference	Assailant and victim '…causing or intending to cause the Victim physical damage…'
Suicide_attack	Cognitive structure Harm reference	Assailant and victim '…causing or intending to cause the Victim physical damage…'
Invade	Cognitive structure Implicit harm	Invader and land '…in an aggressive attempt to cripple or dominate its people'
Fighting_activity	Cognitive structure Violence reference	Combatants '… in a chaotic, violent, and usually protracted Fight…'

As tabulated, the final set of frames, after excluding irrelevant frames and adding conceptually-related ones, is used to construct the macro-cognitive level of the lexicon. These frames represent the raw material which forms the macro structure of the proposed lexicon. They significantly contribute to the construction of the lexicon on the cognitive and linguistic levels. Abuse, terrorism, rape, kidnapping, among others, provide a very poor list of LUs. However, they are essential, on the cognitive level, to crystallizing the frame of violence. They represent a set of typical violent acts. Cause_harm, cause_bodily_harm, killing, and attack, for instance, are evidently rich frames on the linguistic level. They display comprehensive lists of LUs expressing violence.

5.3 Evocative words and Cambridge classification

The extracted list of frames represents more than 150 evocative LUs. Some frames display exhaustive list of LUs, while others record very limited number of LUs. Despite the richness of the definition of 'Violence', only three LUs are provided as evocative of 'Violence': brutality, savagery and violence. The LUs, however, of the 12 frame, including 'Violence', are not linguistically related to each other. More important, they are not linked to the concept of violence in FN, except for 'Cause_harm' LUs.

Therefore, CST is used to look each LU up so that its semantic field or 'topic' can be identified. It exerts some effort to suggest a corpus-based topic for each given sense of a word: LU. For instance, it suggests 'winning and defeating', 'hitting and beating', 'mixing and mixtures', 'shaking and swinging', among others, as candidate corpus-based topics for the word 'beat'. Evidently, only the second topic is selected for being violence-related. 'Cause_harm', the richest violence-related frame, provides LUs varying from the slightest harm, which belongs to Cambridge's hitting and beating topic, to the severest harm, which is categorized under death penalties in Cambridge.

However, the cognition-based definitions of FN LUs and the dictionary definition of different word senses in Cambridge are not always compatible. The more fine-grained senses FN gives to a word, the less effective CST topics are. For instance, 'spear' is listed as a 'Cause_harm' evocative LU in FN. It is defined as a verb denoting 'pierce or strike with a spear or other pointed object'. When searched in CST, spear returns two topics: weapons and hunting and fishing. In this case, the FN definition of the LU is checked to place it within the most similar CST topic. For 'spear', two topics are suggested: hitting and beating and cutting and stabbing.

CST topics, even when they are not directly beneficial, play a vital role in classifying LUs under violence-related categories. They inspire the construction of new violence-specific frames. They also highlight, but informally, the degree of

association between the LU and violence. To elaborate, if a LU is embraced in violence-irrelevant thesaurus topic, its violence-related sense, cited in FN, would not be frequently used. 'Elbow', for instance, is categorized in the body parts topic in Cambridge and in the cause_bodily_harm in FN. Cambridge classification focuses on the most frequent sense of the word, while FN accounts for every potential corpus-based sense. Table 2 exemplifies FrameNet and Cambridge taxonomies of some LUs.

Table 2. LUs frames and their corresponding CST topics

Lexical Unit	Frames	Topics
Jab (v.)	**Cause_harm**	Hitting and beating
	Cause_impact	
Electrocute (v.)	**Cause_harm**	Hitting and beating
Attack (n.)	**Attack**	Physical and sexual assault and
	Offenses	abduction
Bloodshed (n.)	**Killing**	Violent or aggressive

As tabulated, CST can deliver promising results helping in specifying the violent act or show some very general topics linking the act to its superordinate. Still, when CST overgeneralizes the topic of a LU, as in the fourth case, it reiterates its relevance to violence. Thus, CST topics are extensively used in reporting new frames or modifying existing ones.

5.4 LUs and WordNet synsets

After creating all frames and recording their LUs, WordNet 2.1 is enabled to enrich the LU entry with synonyms. It is used to consolidate the integrity of the LUs in a frame and add another hierarchical dimension to the lexicon. WN is based on 'is-kind-of' hierarchical ranking relation. It classifies words in synonym sets (Synsets) and provides various senses to each word sense. It is intentionally selected as it targets word senses: LUs, as FN does. In the assault frame, for instance, the net number of LUs after extracting data from FN and filtering them via CST is 30. After searching these LUs in WordNet, 34 synonymous LUs are introduced. Table 3 displays FN and WordNet definitions of a LU and a list of suggested synonyms.

Table 3. FN definition and WN definition and synsets of a LU

LU	FN definition	WN closest definition	Synsets
Bash (v.)	strike hard and violently	hit hard	sock, bop, whop, whap, bonk
Ambush (v.)	Attack unexpectedly from a concealed position	wait in hiding to attack	Scupper, bushwhack, waylay, lurk, ambuscade, lie in wait
Do in (v)	To kill someone	To get rid of someone by killing	neutralize, neutralise, liquidate, waste, knock off

6.   Results and discussion

Tracing violence-related, whether directly or indirectly, frames results in eight frames. A similarity pattern is concluded after exploring the eight frames. Investigating the pattern in FN data filters the frames and adds new ones. Then, these frames are studied and their LUs are searched for in CST and WN. Finally, a final definition of violence is formed, some frames are modified and others are created.

6.1 Final definition of violence

The lexicon is limited to physical violence. Violence is defined, in the lexicon, as a situation in which either an aggressor commits a physical act to harm a victim or two sides commit physical acts against each other in order to cause harm to each other. So, two situations are cognitively present in the mind of a speaker or a hearer, when violence is mentioned. In the first situation, aggressor and victim are present as core FEs while harm is essentially stated or implied therein. In the second situation, the two sides are not clearly classified into aggressor and victim. The participants mutually try to harm each other. They are either fighter1 and fighter2 or fighters. To elaborate, physical harm is always manifested in the 'Violence' frame. The harm can be caused or intended but it serves a core unexpressed FE. It is either aggressively caused by an assailant against a passive side: victim, or mutually inflicted by two sides against each other in a fight.

The provided definition is slightly from FN's. It elaborates more on the two situations of violence: one-sided and mutual, more explicitly than FN. It also directly links violence to attack and fight. It also excludes the situation related to force and power of nature. So, at the broad level aggressor and fighter FEs exclude each other from the Violence frame. At a more specific level, fighter1 and fighter2

exclude fighters in the mutual-violence-expressing frame. The following examples are cited from the TenTen English corpus and are manually annotated to clarify the two conceptual situations of violence:

(1) Her research interests include [men]^{Aggressor}'s [violence]^{LU} against [women and children]^{Victim}
(2) The state of Guerrero is plagued by drug [violence]^{LU} among [rival gangs] Fighters

6.2 Newly-reported frames

CST topics instruct the process of creating new frames. *Cutting_stabbing* is a new frame embracing LUs from FN's 'Cause_harm' and WN's sense-specific synonyms. It is named after a CST topic. Aggressor or cause, victim, and sharp tool are the core FEs in *Cutting_stabbing*. This frame represents a specific kind of harm which stipulates the existence of a sharp tool, whether linguistically stated or null instantiated. The following examples, representing *Cutting_stabbing*, are cited from the TenTen English corpus and are manually-annotated:

1. [A soldier and a civilian]^{Victim} were [wounded]^{LU} in an attack by [suspected cartel hit men]^{Aggressor} Sharp tool: Null Instantiated
2. [Plath]^{Aggressor} [stabbed]^{LU} [her]^{Victim} with [a knife and a bottle]$^{Sharp\ tool}$
3. [executioner]^{Aggressor}, on seeing this, was ordered to [pierce]^{LU} [him]^{Victim} with [a sword]$^{Sharp\ tool}$

They express a situation in which an aggressor attacks a victim by a sharp tool to cause a physical harm: cut in the body. Although the sharp tool may be null instantiated, it is always cognitively understood in this situation. Thus, it is directly-related to another newly-reported frame: *Sharp_weapon*. *Sharp_weapon* refers to a kind of sharp instrument that is typically used to cause harm by making a cut in the body. The core FEs in *Sharp_weapon* are sharp tool and a carrier or user. Sometimes the carrier is null instantiated. Still, it is always perceived when the frame is evoked. For instance, [They]^{Carrier} carried either short [swords]$^{Sharp\ tool}$ or sticks.

Another new frame inspired by CST and based on 'Cause_harm' LUs is *Hitting_beating*. It refers to a situation in which an aggressor attacks a victim by directing hits in order to cause physical harm. The aggressor either uses one of his/her body parts or an instrument to hit the victim. Moreover, the victim's body part upon which a hit is directed may be linguistically present. Thus, aggressor and victim are core FEs while body part and instrument may be mentioned. The following examples display some incidences of *Hitting_beating*:

1. [The governor]^Aggressor used to [beat]^LU [the citizens]^Victim [with his own hand]^Aggressor's body part

2. [His]^Victim body had [bruises]^LU and signs of torture

<div align="right">Aggressor: Null Instantiated</div>

3. [She]^Aggressor [slaps]^LU [me]^Victim in the [face]^Victim's body part

4. So [he]^Victim was [flogged]^LU with [a knot]^Instrument

<div align="right">Aggressor: Null Instantiated</div>

Very similar to *Hitting_beating* frame is *Breaking*. *Breaking* is also created as a separate frame displaying an aggressor causing physical harm to a victim by breaking the victim's body part. Aggressor and victim's body part are core FEs. It is evoked by sentences like:

(1)   where [soldiers]^Aggressor [smashed]^LU [her head]^Victim's body part

(2)   [Husband]^Aggressor causes [fracture]^LU in [wife's hand]^Victim's body part over huge spending

6.3 Modified frames

Some FN frames are placed in the lexicon with a slight modification to be more violence-specific or to cover a broader violence scope. First, FN's 'Rape' is generalized to be *Sexual_assault* in the lexicon. It is extended to include any physical attack involving sexual offense. *Sexual_assault* indicates a situation in which an aggressor aims at inflicting physical sexual harm against a victim. Thus, more LUs are added to FN's: indecent assault, ravish, ravishment, dishonor, among others. Some of the added LUs, such as dishonor, express a violence-related sense which is contemporary rarely used. Still, the lexicon, aiming at providing a comprehensive coverage of physical violence, includes them. The core FEs in this frame are FN's: perpetrator, victim and the core unexpressed event. Similarly, 'Execution' is enlarged to be *Death_penalty* and more synonyms are added to the existing LUs.

Second, 'Weapons' is changed to include carrier or user as a core FE and to exclude all LUs referring to specific kind of weapons, such as knife, gun, explosive and bomb. Thus, *Weapons* in the lexicon embraces only general LUs summoning weapons to the human cognition, such as arms, armed and weapon. 'Fighting_activity' is transformed into *Fight* and is dissociated from 'Hostile_encounter'. It is linked to the violence frame and modified to have fighter1 and fighter2 or fighters as core FEs. Moreover, it is linguistically enriched by war representative LUs, originally belong to hostile encounter, and WN-driven LUs. For instance, Jihad is driven from WN as a kind of war. 'Attack' and 'Killing' frames are adopted from FN without modification.

Analogues to *Sharp_weapon* is *Explosive_weapon*. It is constructed as a violence-related frame. It includes the special kind of harmful devices which use fires and flames. LUs in *Explosive_weapon* range from guns to bombs. Exploding device and carrier or user are the core FEs. It is a mixture of FN's 'Explosion' and 'Improvised_explosive_device'.

(1) [Four hitmen]^{Carrier} [armed]^{LU} with [machine guns]^{Explosive device} waylaid Shiri

(2) [they]^{User} have not [exploded]^{LU} [bombs]^{Exploding device}

6.4 The lexicon of violence

After adopting, modifying and creating frames, FTF relations are enabled to link the frames and place them into the lexicon. Being related to the same theme: violence, the frames are hypothesized to be cognitively linked. Therefore, the definitions and the FEs of the twelve frames are scrutinized in the light of FN relations. Two relations governing the structure of the frames are detected. First, the inheritance relation is dominatingly present among the frames. Violence is inherited by both Attack and Fight. Each child frame represents one of the violence scenes. Attack realizes the aggressor/victim situation, while *Fight* embodies the aggressors or fighters scene. Still, the intended, expressed or unexpressed, physical harm is present. Moreover, Attack is inherited by a number of harm-specific frames: *Hitting_beating*; *Cutting_stabbing*; *Breaking*; *Sexual_assault* and Killing which is inherited by *Death_penalty*. The frame of *Weapons* is a parent inherited by *Sharp_weapon* and *Explosive_weapon*. Second, use relation is saliently manifested among some frames. It explains links *Weapons* to *Violence*, *Attack* and *Fight*. It also links the more weapon-specific frame: *Sharp_weapon* to the more attack-specific frame: *Cutting_stabbing*.

7. Conclusion

This study applies the cognitive linguistic approach of FN to build a thematic monolingual dictionary tackling the multidisciplinary concept of violence. It adopts FN methodology and integrates FN data, CST topics and WN synsets into a cognition-based lexicon. The created lexicon is structured after FN. It divides the concept into two scenes: one-sided and mutual violence, under which a number of frames are included. It consists of twelve frames, some of them are adopted from FN, some are modified and others are newly-created. The lexicon is cognitively-integrated by two FN relations: inheritance and use. English TenTen corpus is used to validate the results of the lexicon.

The study uses the perception of violence, as reflected in language, as a key to defining the debatable concept: violence. The three recruited resources provide compatible, to some extent, information. Despite following different categorization methods, they display very similar definitions of LUs. Placing the same LU in a corpus-based concluded topic, cognition-based frame and synonymy set suggests new linguistic and cognitive insights into the multidisciplinary concept exploration. Future work in this area may move to a more abstract level covering verbal or psychological violence. Expanding the lexicon to other languages should highlight cognitive and linguistic similarities and differences among languages.

Acknowledgements

References

Atkins, S., Rundell, M., & Sato, H. (2003). The contribution of FrameNet to practical lexicography. *International Journal of Lexicography*, 16(3), 333–357.

Atkins, B., & Rundell, M. (2008). *The Oxford guide to practical lexicography*. Oxford University Press.

Baker, C. F., Fillmore, C. J., & Cronin, B. (2003). The structure of the FrameNet database. *International Journal of Lexicography*, 16(3), 281–296. doi: 10.1093/ijl/16.3.281

Baker, C., Fillmore, C., & Lowe, J., (1998). The berkeley framenet project. In *Proceedings of the 36th Annual Meeting of the Association for Computational Linguistics and 17th International Conference on Computational Linguistics-Volume 1*, (pp. 86–90). Association for Computational Linguistics. doi: 10.3115/980845.980860

Bertoldi, A., & de Oliveira Chishman, R. L. (2011). Developing a frame-based lexicon for the Brazilian legal language: the case of the criminal_process frame. In *International Workshop on AI Approaches to the Complexity of Legal Systems*, 256–270.

De Cao, D., Croce, D., & Basili, R. (2010). Extensive Evaluation of a FrameNet-WordNet mapping resource. *Proceedings of the Seventh conference on International Language Resources and Evaluation*, Malta: ELRA.

Evans, V. & Green, M., (2006). *Cognitive Linguistics: An Introduction*. Edinburgh: Edinburgh University Press.

Fillmore, C. (1975) 'An alternative to checklist theories of meaning', *Proceedings of the First Annual Meeting of the Berkeley Linguistics Society*. Amsterdam: North Holland, 123–31.

Fillmore, C. and Baker, C. (2010). A frames approach to semantic analysis. In B. Heine and H. Narrog, (Eds), *The Oxford Handbook of Linguistic Analysis*, (pp. 313–340). Oxford: Oxford University Press.

Fillmore, C. J., Johnson, C. R., & Petruck, M. R. (2003a). Background to framenet. *International journal of lexicography*, 16(3), 235–250. doi: 10.1093/ijl/16.3.235

Fillmore, C., Petruck, M., Ruppenhofer, J., & Wright, A. (2003). FrameNet in action: The case of attaching. *International Journal of Lexicography*, 16(3), 297–332 doi: 10.1093/ijl/16.3.297

Ostermann, C. (2012). Cognitive lexicography of emotion terms. *Proceedings of the 15th EURALEX International Congress*, 493–501.

Ostermann, C. (2015). *Cognitive lexicography: A new approach to lexicography making use of cognitive semantics*. Berlin, Boston: De Gruyter Mouton doi: 10.1515/9783110424164

Ruppenhofer, J., Ellsworth, M., Petruck, M. R., Johnson, C. R., & Scheffczyk, J. (2016). Framenet II: Extended theory and practice. URL http://framenet.icsi.berkeley.edu/book/book.pdf.

Schmidt, T. (2007). The kicktionary: A multilingual resource of the language of football. *Data Structures for Linguistic Resources and Applications. Gunter Narr, Tübingen, Germany*, 189–196.

Schmidt, T. (2009). The Kicktionary–A Multilingual Lexical Resource of Football Language. In H. Boas (Ed.), *Multilingual FrameNets: Methods and Applications*. (245–286). Berlin: De Gruyter.

Subirats, C., & Petruck, M. (2003). Surprise: Spanish FrameNet. In *Proceedings of the Workshop on Frame Semantics at the XVII. International Congress of Linguists*, Prague.

Author's address

Esra' Moustafa Abdelzaher
Department of English Language, Faculty of Al-Alsun
Ain Shams University
Cairo
Egypt

esraa.abdelzaher@alsun.asu.edu.eg