

Thematic Roles: Primitives of Semantic Theory or Inferences from LCS

DENISA ELENA DRĂGUȘIN
PhD Lecturer
Faculty of Letters
Spiru Haret University, Bucharest
Romania

Abstract:

The paper proposes an analysis of thematic roles concepts. Following Fillmore's A Case for Case thesis, it explains what theta roles are, how they are structured or ranked and the manner in which they are integrated into syntax. It also considers a list of thematic roles that have been found by researchers to be relevant in verb classification. The aim of the paper is to analyze and explain notions like case-frame, thematic relations, thematic grid, lexical conceptual structure a.s.o. The analysis of thematic roles leads us to the conclusions that thematic roles are not primitives of semantic theory, but they are inferred from the Lexical Conceptual Structure (LCS) of the predicates, and that there is no direct (systematic) correspondence between them and morphological cases or syntactic functions.

Key words: thematic role, argument structure, inherent/non-inherent arguments, morphological case

1. Fillmore's Case Grammar

In 1965 Noam Chomsky published his linguistic study named *Standard Model of Generative Transformational Grammar*.

Later on this theory will be labelled as Universal Grammar. Chomsky's main idea is that every sentence in any

language has two levels of representation: a deep structure and a surface structure.

Later, Charles Fillmore who was proponent of Chomsky's theories wrote in 1968 his study named *The Case for Case* which was a more in depth analysis about the fundamental importance of semantics in syntactic and morphological phenomena or how semantics triggers changes in syntactic and morphological representation¹. Fillmore's theory was that every verb (predicate) selects different deep cases. The combination of cases that a given predicate could have is called **case frame** or **thematic grid** or **argument structure**. In this paper we will refer to this combination as **case frame**.

We will see that although the theory of **case grammar**² or **thematic relations** is an universal semantic theory, there are many differences between all over the world languages:

A common assumption is that the universal base specifies the needed syntactic relations, but the assignment of sequential order to the constituents of base structures is language specific. (Fillmore, *The Case for Case*, 1968).

For a better understanding of Fillmore's theoretical intercession, first we must comprehend the difference between deep and surface structure³. Secondly we also must understand that in **case-grammar**, thematic roles are concepts which only express relation between the participants (arguments) in an event (denoted by the verb). θ -roles are descriptive labels which do not have theoretical status: θ -roles are *defined* notions.

¹ Fillmore's fundamental theory was that grammatical functions as subject or object can be determined by the deep semantic valence of the verb, which is "projected" into its surface structure grammatical categories as Subject and Object and in grammatical cases such as Nominative and Accusative.

² The case grammar is Fillmore's semantic theory that deals with predicates' structure of events and their participant structure.

³ If there is more than one case form that appears in the surface structure of the same sentence, it means that either more than one deep-structure case is involved or the sentence is complex.

A good example regarding Fillmore's statement is the difference between *pro-drop parameter* and *non-drop parameter* languages: i.e. the difference between English and Romanian. For example let us put in discussion the transitive verb *to leave* (*a părăsi*):

1(a) *I left the house.*

1(b) *Am părăsit casa.*

The deep-structure subject appears in both examples, but as we can see there are important differences between English and Romanian regarding surface structure subject. In 1(a) the Subject 'I' is obligatorily present, otherwise the sentence would be ill-formed. In 1(b) the Subject does not appear as an "independent" lexical item 'Eu', but it is expressed by the inflectional marker that forms the past tense (Perfect Compus) of the Predicate. This is a short example that underlines the deep and surface structure's major distinction.

2. Case-structure. Thematic Roles

We will take a closer look to case-structure by taking into account the most relevant cases (thematic roles) which according to Fillmore are concepts implied in identifying certain types of judgments that human beings are making about the events that are taking place around them⁴. The following list

⁴ In his 1968 *The Case for Case* thesis, Fillmore identified the following "appear to be needed" cases: *Agentive (A)*, is the case of animate instigator of the action of the verb; *Instrumental (I)*, is the case of inanimate object (the causer) involved in the action identified by the verb; *Dative (D)*, the case of animate being affected by the verb; *Factitive (F)*, the case of object or being resulting from the action identified by the verb; *Locative (L)*, the case of location or spatial orientation identified by the action of the verb; *Objective (O)*, the most neutral case, from semantic point of view, the case of anything representable by a noun whose role in the action is identified by the verb itself. Fillmore has come to the conclusion that more cases are needed in order to express a more in depth semantic differences between roles which have been found to be relevant in verb classification (see above).

comprises the more relevant thematic roles which are universally assigned by the verb:

Agent (A) is the animate initiator, the doer of an action and is usually responsible for the action. The subject occurred in the specifier position of VP is an Agent:

- a. *Mary makes a cake.*
- b. *John is swimming.*
- c. *Sam killed a cockroach*

Theme (T) was introduced by Gruber (1965) and is the least consistently used term of all thematic relations (a default thematic role) and describes the entity that undergoes a change or the entity which is perceived. It occurs only with verbs of motion or location (both can be either concrete or abstract):

- a. *The book is on the shelf.*
- b. *She threw the ball.*

Patient and **Percept** are more specialized roles that have been created for such “subtypes” of themes, in order to capture the semantic differences between them:

- a. *Joan ate an apple.- Patient, the entity which undergoes a change or suffers an action*
- b. *I hit the target.- Patient*
- c. *The shadow scared Mary.- Percept, designates the entity which is perceived or experienced*

Experiencer (E) is the role of the animate being affected by the action or state identified by the verb; the animate participant experiencing a psychological process:

- a. *Joan dislikes her sister.*
- b. *The lion frightened the tourists.*

Benefactive is the entity that benefits from the action denoted by the verb:

- a. *I have made **him** a favour.*
- b. *Mike built a house **for his father**.*

Goal (G) represents the location or entity towards someone or something moves:

- a. *I am going **to Paris**.*
- b. *I gave **Mary** a book.*

Source (S) is defined as the entity from which someone or something moves:

- a. *John left **his home**.*
- b. *She fell **from the cradle**.*

Instrument (I) designates an object that helps performing an action:

- a. *They broke into the house **with a false key**.*
- b. *My mother's **sharp knife** cut me to bone.*

Location (L) is the specification where something takes place:

- a. *I saw him **at the theatre**.*
- b. *The fans are celebrating **on the streets**.*

Path is the trajectory which an entity covers:

- a. *He is swimming **towards the current**.*
- b. *She walks **along the edge**.*

Thematic roles acquire substance only in the context of the verbs (predicates) that require them, in other words, they are not inherent, but relational concepts. θ -roles are read off from the LCS (Lexical Conceptual Structure) of the verbs.

Since cases are “linked” to the arguments of the verbs, we must emphasize on the fact that there is no systematic relation between roles and morphological cases or syntactic functions: i.e. the Nominative case, and consequently the Subject can be expressed by the θ -role of Agent (*John is running.*) or Experiencer (*Mary loves John.*) or Theme (*Jane*

fell.) a.s.o. Thematic roles can also appear in more than one syntactic functions: i.e. Theme can be Nominative Subject (*The window opened.*) or Accusative Direct Object (*He broke the window.*) a.s.o.

Furthermore, Fillmore advanced the idea that some cases are more prominent than others. In his attempts to relate argument structures and syntax in order to establish the principles for selection of Subject and Direct Object, Fillmore has come to the conclusion that cases associated with a verb, do not form an unordered set of roles but they are ordered according to the *relative degree of prominence* principle: i.e. for the change of state verbs Fillmore has developed, based on his thematic hierarchy⁵, the following subject selection principle:

If there is an A [=Agent], it becomes the subject; otherwise, if there is an I [=Instrument], it becomes the subject; otherwise, the subject is the O [=Objective, i.e., Theme/Patient]. (Fillmore 1968:33)⁶.

The examples above raise another problem, namely the co-occurrence restrictions (see examples *a* and *b*). In example *a*, the Agent is the subject, while in example *b*, the Instrument is the subject. The subjects of *a* and *b* are different at grammatical level as well as at the lexical features (i.e. the Agent in example *a* is [+animate] while in example *b* is [-animate]) and that

⁵ Thematic hierarchy is a language-independent ranking of possible semantic roles, which establishes prominence relations among them with respect to argument realization: i.e., subject and object selection.

⁶ Let us take for example the verb *to break* which has the following lexical entry: (Agent) > Theme/Patient (Instrument):

Mary broke the window with a hammer.

- a. Mary broke the window. Agent is the S-structure Subject*
- b. The hammer broke the window. Instrument is the S-structure Subject.*
- c. The window broke. Theme/Patient is the S-structure Subject.*

Giving the examples above we may conclude that the verb *to break* needs at least one argument bearing the role of Theme/Patient in order to form a grammatically correct sentence. The thematic roles of Agent and Instrument are optional. One may also conclude that the argument of a verb bearing the highest-ranked semantic role is its Subject.

explains the fact that the combined meaning of the two sentences is not produced by conjoining their subject. Thus the following sentence is ill formed: *Mary and a hammer broke the window*. We can conclude that only noun phrases representing the same case and lexical features may be conjoined: i.e. *Mary broke the window; John broke the window; Mary and John broke the window*.

Fillmore very well noticed, regarding the Locative case, that are superficial, yet important differences between locational and directional elements (role-concepts) determined by the character of the associated verb. Thus, the thematic domain of movement and location has emerged⁷.

It is well known that concepts are structured in semantic fields or *frames*, thus role-concepts come from several different semantic frames such as the field of movement and location (Source, Goal, Location, Theme, Path, etc.), filed of human action and causation (Agent, Theme/Patient, Benefactive, Cause, etc.) and each concept may have complex structure, with interrelating subcomponents which may belong to different frames (fields): i.e. consider the unergative verb of movement *to run (a alerga)*.

- a. *She is running to the park.*
- b. *She is running in the park.*
- c. *She is running away from him.*
- d. *She is running for him.*

In examples *a, b and c*, all three adjuncts⁸ (*to the park, in the park and away from him*) belong to the same semantic field of

⁷ Jackendoff (*Consciousness and the Computational Mind*, 1987) has studied categorization and concept formation, that is the way concepts are stored in the mental lexicon, so that inferential relations between concepts and lexical items that express these concepts, are made available and how lexical items are associated with LCS (Lexical Conceptual Structures).

⁸ Adjuncts are Prepositional Phrases which are not sisters to a prepositional verb. They may occur as optional (an excessive argument), being adjoined to the verb. Their function is either: *a. free Prepositional Object* (He traveled

movement and location, bearing the Goal theta role in *a*, the Location theta role in *b* and Source theta role in *c*. In example *d* the adjunct *for him* no longer belongs to the semantic field of movement but to the semantic field of human action and causation, bearing the Benefactive theta role.

Since morphologic case and syntactic function seldom identify a semantic role, the more explicit markers of roles are prepositions. Hence, for Goal we have preposition *to*, for Location preposition *in*, for Source preposition *from*.

Jackendoff has observed that Source-Goal cluster defines a Path which can be divided in three subtypes: a. *bounded paths*, when the reference point (Goal) is reached: i.e. *She arrived home*; b. *directions*, when the reference object does not fall on the path: i.e. *She is walking **towards** her home*. (with Goal directional preposition); *Mike walks **away from** his car*. (with Source directional preposition); c. *routes*, when the reference object or place is merely related to some point in the path: i.e. *They are walking **along** the sidewalk*.

As we can see above there are typical (specialized) prepositions that define paths like: *towards*, *along*, *across* etc., but there are also verbs that have an inherent Source, Goal or Direction meaning, which are more limited in their distribution, verbs like: *rise/fall*, *ascend/descend*.

The verbs *rise/fall* imply **unintentional** (involuntary) vertical movement (upwards or downwards), and the direction is either explicitly mentioned (*The sun rose **up** from the sea. / The pencil fall **down** to the ground.*) or incorporated as a part of the Source or Goal constituent (*The sun **is rising**. / The rock **is falling***). Notice that *rise* signifies upward movement only (*The sun rose down* is ill formed), while *fall* signifies downward movement (*The pencil is falling up* is ill formed also)⁹.

with his girlfriend.) or b. Adverbial Modifier (He is talking *through a speaking-tube*.)

⁹ There is a strong link between subcategorization frames (transitives, unergatives, unaccusatives), thematic roles and the intentional/unintentional aspect.

The verbs *ascend/descend* imply **intentional** (voluntary) vertical movement. They also signify movement along vertical dimension with the difference they, unlike *rise/fall* cluster, always incorporate the *up/down* directional particles: i.e. *Mary is ascending the stairs.* / * *Mary is ascending **up** the stairs;* *He is descending the ladder.* / * *He is descending **down** the ladder.*

The second conceptual structure, as it was mentioned before, is the field of action and causation (Agent, Patient/Theme, Cause, Instrument). Jackendoff concluded that the generic verb in the action field is ACT or DO¹⁰, involving an Actor or a Doer (Agent or Instrument/Cause) and an optional Theme/Patient. He also distinguished between causative [+agency] and non-causative [-agency] verbs of movement:

- a. *Mary moved the chair. (causative)*
- b. *The chair moved. (non-causative)*
- c. *The storm rolled the ship. (causative)*
- d. *The ship rolled down the wave. (non-causative)*

It can be clearly noticed that there is a direct relation between non-causation and intransitivity (unaccusatives). Whenever the verb is used transitively, the Subject bears the theta role of Agent or Instrument/Cause and when is used intransitively, the Subject bears the theta role of Theme/Patient.

Grubber (1965) identifies a second kind of agency called *permissive agency*, which does not obstruct the movement: i.e. *Mary let them eat the cake.*

Fillmore (*Current Issues in Linguistic Theory, 1977*) also suggests, regarding relative prominence relations among

¹⁰ David Dowty, *Thematic Proto-Roles and Argument Selection, 1991*, developed the concept of proto-roles which decomposes semantic roles into more basic components, but these components do not constitute a set of necessary and sufficient conditions on any given role. Hence, the Agent Proto-Role will have the following contributing properties: *volitional involvement in the event or state; causing an event or change of state in another participant; sentience and/or perception; movement (relative to the position of another participant)*

semantic entities, that Thematic Hierarchy could be replaced with a series of semantic rankings as follows:

- a. An active element outranks an inactive one;
- b. A causal element outranks a non-causal one;
- c. A human (or animate) experiencer outranks other elements;
- d. A changed element outranks a non-changed one;
- e. A complete or individuated element outranks a part of an element;
- f. A “figure” outranks a “ground”;
- g. A “definite” element outranks an “indefinite” one.

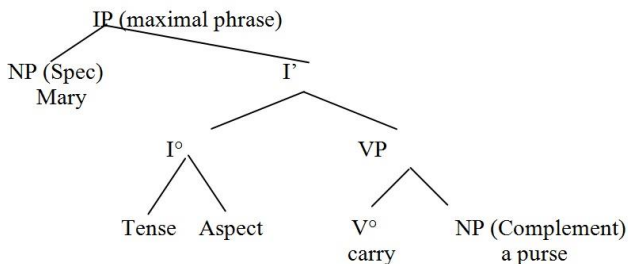
3. Integrating Theta Roles into Morphology and Syntax

Fillmore’s case frame theory (thematic theory) raises a major problem, namely how to systematically integrate (relate) theta roles into morphology and syntax. His early attempts were based on the theory that case-frames are projected as deep structures and derived to surface structures through rather an arbitrary transformation.

Every predicate theta marks all the syntactic positions for which it is subcategorized and the roles in the predicate’s theta grid match the subcategorization frame: i.e. the verb *to open*, when used transitively (*I/The wind opened the door.*), has the following theta grid (argument structure) and subcategorization frame: <Agent/Instrument, Theme/Patient>/[_NP]. As we can see, *any syntactic position capable to bear a theta role is an Argument, and any NP filling an argument position is an argument, therefore the subject and the object are arguments of the verb.* Another consequence emerges, that is one and the same role cannot be assigned to two different arguments (*I opened **the door, the window.***), neither is possible for the verb to assign two thematic roles to the same argument (***She** admired*). This is called the **theta-criterion** theory.

Having said all that, it is necessary to discuss and explain the differences between the *external* and the *internal* arguments of the verb. The *internal argument* is the one which is dominated (receives its role) by the first projection of the head: i.e. *Mary is caring a purse*. A **purse** is in Complement position (post-head sister of V°) and receives its role from the verb head; in most cases corresponds to the Patient/Theme role¹¹.

The *external* argument position (in the example above *Mary*, which is the Agent) is occupied by the Subject which is base generated in the Specifier position of the whole VP (the maximal V-projection)¹² which c-commands it, as in the example below:



Back to *internal* arguments, a predicate can theta mark only one argument directly, which can function as Direct Object of transitive verbs and is called *direct internal* argument: i.e. *Anna loves red roses*. (*Ana iubestetrandafirii rosii*). Other internal arguments (Prepositional Phrases) receive their theta

¹¹ Some theta roles are realized as compulsory constituents of the whole predication underlying a sentence. These roles assigned by the verb to its arguments are called *inherent* theta roles: Agent, Experiencer, Patient/Theme. Other roles assigned by the verb to its arguments like Location are realized by Adverbials. These roles are called *non-inherent* theta roles.

¹² The **VP- internal hypothesis** is the theory according to which the Subject is base-generated inside the VP, in the position of Specifier. From this position, it will have to move to the higher Specifier position under the IP node. This movement is **obligatory** because English grammar, as we have mentioned, does not allow null Subject at the sentence level that is, all sentences must have an expressed Subject in Specifier of IP (Inflectional phrase) position at Surface structure level.

roles from the predicate indirectly by the joint contribution of the verb and its preposition, their syntactic function being of a Prepositional Object. They are called *indirect internal* arguments: i.e. *She looked for **her wallet***. (*Ea își căuta portofelul*). As we can see, from the indirect internal example, there is a big difference in regard with how the verb assigns morphological case and syntactic function to its argument. While in both English and Romanian the morphological case is the same (Accusative), the syntactic function differs completely. In English the argument of the verb is a Prepositional Object (indirect internal argument), while in Romanian the argument of the verb is Direct Object (direct internal argument). That is because, *looked for* is a phrasal verb, while *căuta* is a transitive verb.

It can also be noticed that in both examples, the verbs assign the same Patient/Theme theta role to their arguments. This leads us to the conclusion, as we have said before, that thematic roles are not primitives of semantic theory, but they are inferred from the Lexical Conceptual Structure (LCS) of the predicates, and that there is no direct (systematic) correspondence between them and morphological cases or syntactic functions.

Lexical Conceptual Structure (LCS) *is a compositional abstraction with language-independent properties that transcend structural idiosyncrasies* (Jackendoff, 1983). In other words, LCS is a semantic level of representation that decomposes certain aspects of the meaning of the verb into more elementary predicates (DO, CAUSE, GO, BE) and links them to predicate's argument structure: i.e. the verb *give* has the following LCS: $[[x \text{ does something}] \text{ cause } y \text{ [to be at } z]]$ (*John gave a book to Mary*). Since the theta roles provide only a partial representation of the meaning of the verb, LCS offers a more complex representation of predicate's meaning. The relation between LCS and syntax is mediated by case-frames

which contain only indications of relative prominence between arguments (variables)¹³.

Let us make a more in depth analysis of the statements above. Consider the following sentence:

a. John filled water into the jar.

The case frame would be as follows: <Agent, Patient/Theme or Locatum, Goal>. But we can also say:

b. John filled the jar with water.

The sentence will be grammatically correct also with the following case-frame: <Agent, Patient/Theme, Locatum or *Displaced Theme*>. The question is how we choose between these two sentences, what is the correct case-frame “matrix” of the verb *to fill*. First of all, the Lexical Conceptual Structures of the two sentences are as follows: *a. [x does something that cause y to be at z]*; here *z* is seen as a Goal so sentence *a.* has a Locative tinge ($x < y, \text{loc } z >$). But the verb *to fill* is rather a change of state verb then a verb of movement, so the LCS of example *b.* looks like this: *[x does something that cause z to come into a STATE by means of causing y to come to be at z]*; here *z* is seen as a Patient/Theme that suffers a change of state from the verb directly ($x < z, \text{with } y >$), being the direct argument variable. Secondly, the *x* and *z* variables in the second example are more prominent than *y*¹⁴, which occurs only in the **means** clause so that the case-frame “matrix” of the verb *to fill* looks as follows: <Agent, Patient/Theme, Locatum>. This is an example how Lexical Conceptual Structure of a predicate and the theta grid or argument structure of a predicate work together to the well-formedness of a sentence.

¹³ The argument places of predicative constituents (external or internal arguments, direct or indirect arguments) are held by variables which are place holders for NP.

¹⁴ Fillmore (*Current Issues in Linguistic Theory, 1977*) has come to the conclusion that roles are ranked not in terms of their position in an event, but rather in terms of the semantic components that contribute to prominence or perhaps “topic worthiness”.

4. Conclusions

Although we have tried to present the theories in regard with theta roles, what they are and how they are projected into syntax, still many questions and problems remain unanswered. This is because, the Universal Grammar is only a model to follow, with *universally shared properties* of natural languages which form invariant systems and *particular properties* of natural languages representing cross-linguistic differences, possibly described as *variables*. The very variables make so difficult to implement these theories as a very general principle.

Even though there have been many debates among researchers about how a thematic hierarchy must look like (see Chomsky, Grimshaw, Jackendoff, Dowty, Fillmore's thematic hierarchy, ranking), and many changes have been made over years to their theoretical concepts, one thing is clear; that thematic roles are convenient logico-semantic concepts that have a major importance to the syntax of all languages. These differences should be determinable from what each ranking aspires to represent and what generalization it is intended to capture. A particular *thematic hierarchy* may remain a convenient way of stating a valid generalization even though it is derivative. The examination of the existing *thematic hierarchies* provides a window into the nature of semantic prominence relations, suggesting that some derive from the event structure, while others derive from finer properties of the event.

REFERENCES

- Avram, L., 2003. *English Syntax. The Structure of Root Clauses*. Oscar Print, București
- Chomsky, N., 1981. *Lectures on Government and Biding: the Pisa Lectures*. Floris

- Cornilescu, A., 1995. *Concepts of Modern Grammar*. EUB
- Fillmore, C., 1968. "The Case for Case", *Universal in Linguistic Theory*. New York
- Grimshaw, J., 1990. *Argument Structure*. MIT Press
- Jackendoff, R., 1983. *Semantics and Cognition*. MIT Press
- Levin, B., 2005. *Semantic Prominence and Argument Realization II*. MIT
- Massam, D., 1979. *Case Theory and the Projection Principle*. York University
- Stalmaszczyk, P. *Theta Roles and the Theory of Theta-Binding*. University of Lodz
- Șerban, D., 2006. *The Syntax of English Predications*. România de Mâine