Towards resolving two fundamental issues
in semantic role theory

Aliaksandr Huminski

Institute of High Performance Computing

Department of the Social & Cognitive Computing

Singapore

guminsky@hotmail.com

To my teacher Victor Martynov

Abstract

The current state of semantic role theory is investigated to understand the reason or reasons why, even after so long, there remains no agreement to date about two fundamental issues:

- how many roles are needed and what they are;
- how to distinguish a valid set of roles from an invalid one.

The goal of this paper is to explore possible answers, not to assume they are final.

Table of Contents

1. Introduction

The notion of a semantic role[1] is used as a way to capture semantic commonality between events represented by sentences. It was a kind of natural language generalization to extract semantic invariant from a variety of syntactic representations. More than fifty years ago semantic roles were introduced by J. Gruber (1965) and Ch. Fillmore (1968). At that time, six roles were proposed by Fillmore in his seminal article "The case for case"[2]. Later the update included eight roles (Fillmore, 1971).

Afterwards, semantic roles were developed in voluminous literature with enormous proposals regarding the nature of roles and their inventory needed for coverage. Nevertheless, there is still no consensus to date about two of the most important issues.

Role inventory: how many semantic roles are needed and what they are[3]. The range varied from a few (Anderson, 1971) to hundreds of roles (Pollard and Sag, 1994; Fillmore et al., 2002).

Role combination: how to distinguish a valid set of roles from an invalid one. It is unclear why some sets of roles represent classes of verbs, while others do not represent any verb at all.

These two major issues[4] and a kind of uncertainty in the selection of a proper approach to resolve them decreases the overall explanatory effectiveness of semantic role theory and

---

[1] The following notions: participant role, thematic role, thematic relation, theta role, θ-role, deep case, deep semantic case, and semantic case will be considered hereafter as synonyms of a semantic role.

[2] Case Grammar was originally motivated in part by the view that the various possible syntactic configurations in which a verb's arguments could appear (e.g. *John opened the door with a key, The key opened the door,* and *The door was opened by John with a key*, but not **The key opened the door by John*) were more systematically describable in terms of semantic Deep Cases (which corresponded to Gruber's thematic relations) than in terms of the deep structures and transformations of Chomsky (Dowty, 1991:562).

[3] One of the multiple opinions about that is the following: "there is no consensus among linguists on the importance and contents of θ-roles and some researchers seem to diverge from explicit reference to θ-role labels. This tendency is motivated by the fact that there appear not to exist any clear criteria for determining what θ-roles given arguments bear" (Stalmaszczyk, 2007:99).

causes serious drawbacks to the current state of art in semantic role theory and various applications in QA-systems, machine translation and information retrieval.

In linguistics, it even led to negative opinions of a semantic role. They varied from criticism of the entire construct -- there is no construct as murky in any subdivision of linguistic theory as that of 'thematic role' (Newmeyer, 2010) -- and doubt as to whether such a notion exists (Carlson, 1998), to the non-optimistic generalization that semantic roles have not lived up to their initial promise (Levin, 2014). The two issues mentioned are rather complicated to be resolved and have a direct connection to the construction of an adequate theory of meaning[5].

Why did this happen? A huge disproportion in semantic role lists is just a consequence of disparate presuppositions regarding the nature of a role or "lack of a principled basis for determining the semantic content of case roles, and thus for identifying the roles played particular arguments of clauses" (DeLancey, 1991:338). They are mostly justified by intuition and the only evaluation methods have been inter-annotator agreement (Allen and Teng, 2018). In short, roles are poorly defined.

It is uncertain what kind of data motivates a thematic role type. Conceivably, the difficulty we have had in reaching agreement on just what a theory of thematic roles should look like is analogous to that of the blind men examining the elephant, each touching a different part of its body (Dowty, 1991).

A lack of criteria leads to terminological vagueness and inconsistency, allowing for various interpretations. For example, a definition can imply multiple levels of specification (granularity) for semantic roles and the answer to the original question -- *what is the list of semantic roles* -- depends on it, in the same way as what can be seen on a map depends on its scale[6]. As a direct consequence, there is confusion among different approaches, which in turn triggers follow-up discrepancies in role extraction and evaluation.

To the best of our knowledge, the ways in which to resolve the issues are rather general, and did not go beyond statements about the necessity of deeper analysis and objective criteria for role construction[7].

We paid attention to the following gap. The definitions of a semantic role are no doubt a core element in semantic role theory. Nevertheless, despite the fact that there is "a notable absence in consensus about what thematic roles are" (Dowty, 1991), it is hard to make a

---

[4] In fact, there is one more issue formulated in (Dowty, 1989:70): the lack of any effective way to independently justify the assignment of noun phrases to thematic roles in particular sentences. But this issue assumes as a prerequisite that semantic roles are applied to the syntactic level of a language which makes this issue not the main concern.

[5] Theory of meaning is considered as one of the 6 challenges in linguistics (Jackendoff, 2007).

[6] There are three types of resources depending on the level of role granularity. The first level is very specific with roles like *eater* for the verb *eat* or *hitter* for the verb *hit*. In contraposition to it, the third level is very general with the range of two to nine roles. The second level is located between them and contains approximately 10 to 50 roles (Huminski et al., 2019).

[7] For example, in (Petukhova and Bunt, 2008) it was stated that semantic roles describe the way the participant is involved in an event, rather than by internal properties (e.g. does it act intentionally, is it affected, changed, manipulated by the other participants in an event, does it come into existence through the event, etc.). Meanwhile, the reason as to why there is no agreement on the question of what an agent really is lies in the difficulty of characterizing what sentience, volition, intention, causation and even animacy, actually are. These notions are barely, if at all, clearer than the concept of agentivity (Kasper, 2010:8).

comparative analysis of the definitions. Some of the definitions are too general, some look ambiguous and almost all of them are defined in an informal manner, using some examples but without any mathematical concepts or language[8]. Inevitably, when terminology is not strictly defined, it makes any comparison difficult.

So, the analysis of semantic role definitions will be the first necessary step to shed light on the original issues and to investigate why they still exist.

This paper is organized as follows. We are going to start (ch.2) with different approaches to semantic roles and collections of similar definitions with their follow-up transformations into "more mathematical (formal)" representations. In ch.3, analyses of two issues will be made based on pre-processing of the definitions. In ch.4, a new approach will be provided to offer answers on the issues under investigation. Ch.5 will conclude the article. In appendix (ch.6) a comparison of semantic roles is provided.

2. Formalization of semantic role definitions

2.1. Procedure of analysis

A semantic role has been defined in numerous ways, making the notion rather vague. The definitions vary and differ drastically. As we will see, the same phrase, "semantic role", is attached to different concepts, due to the various approaches towards understanding semantic roles. It makes sense to come back to the statement that a semantic role needs revision: "one thing that is clear is a need for a more precise definition of a thematic role in an event" (Carlson, 1998).

The importance of a certain, mathematically formal, and unambiguous definition is no better illustrated than in the particular case of semantic role theory, where ill-defined semantic roles inevitably lead to issues. The approach we are going to follow is to make the definitions of a semantic role as transparent as possible to reach a clear understanding. They will be analyzed and evaluated to some extent, and this analysis will be multifactorial. The main factor is the level of a language to which a semantic role is applied. Based on the level, similar definitions will be grouped into clusters. If there are two or more different approaches inside the same level, sub-clusters will be extracted.

For each cluster (or sub-cluster), the following sequence of information will be provided:

- Examples of actual definitions suitable for the cluster;
- Specific semantic roles extracted on the grounds of the definition;
- Formal definition for the cluster;

---

[8] We guess, the closest project to ours is TRIPS with an attempt of role axiomatization (Allen and Teng, 2018).

- Range of variables in formal definition[9];
- Analysis (of variables and the definition itself).

2.2. Semantic role applied beyond any level of a language

In this subchapter, we accumulate definitions of a semantic role that are not related to any level of a language.

Examples of actual definitions suitable for the cluster.

A semantic role is the role played by a participant in a situation (Begum and Sultana, 2013).

Semantic role is the actual role a participant plays in some situation and it always depends on the nature of that situation (Phuong, 2017).

Semantic roles are symbolic entities that describe the function of the participants in an event from the point of view of the situation in the real world (Morante, 2008).

Semantic roles are the different functions participants have in events (Kittila et al., 2011).

Semantic roles are defined as relational notions which link a participant to some real or imagined situation ('event') (Petukhova and Bunt 2008:41).

Specific semantic roles extracted on the grounds of the definition. The varieties of empirically extracted semantic roles from the definition do not differ from the varieties related to the classical approach, where a semantic role is considered a relation between an argument and a verb in the sentence (see subchapter 2.3.).

Formal definition for the cluster. According to the cluster, a semantic role is defined apart from the linguistic encoding, i.e. purely external. A semantic role ($SR_{P_i}^{E_j}$) refers to a role (RL) of a participant (P_i) in the event (E_j).

$$(1) \qquad SR_{P_i}^{E_j} = RL\left(P_i, E_j\right)$$

Range of variables in formal definition. Participants and an event are explicit variables, but there is no indication if the definition (1) is valid for every participant P_i and every event E_j involved. In other words, it is uncertain if the following formal representation is valid:

$$SR_{P_i}^{E_j} = RL\left(P_i, E_j\right) \qquad (i = \overrightarrow{1, n_j}) \ \& \ (E_j \in \{E\})$$

where a participant P_i runs through all n_j participants in the event E_j and the event E_j runs through the set of all events $\{E\}$.

---

[9] Range of variables involved in a formal definition is important since overstepping the limit might create a contradiction. For example, if the issue with role inventory is formulated as *what is the list of semantic roles to be applied across all verbs to all syntactic constituents?* there would be a contradiction in the formulation if at least one of following cases occurs:

a) semantic roles cannot be applied to all verbs OR;

b) semantic roles cannot be applied to all syntactic constituents OR;

c) semantic roles cannot be applied to the syntactic level of a language.

Analysis. First, there is a redundancy of concepts in (1). A new concept of a semantic role (SR) is defined through the no more understandable concept of a participant role (RL) in the event.

Second, definition (1) operates with the notions of a participant and an event which are not defined in advance. Definitions of an event in the different resources do not provide sufficient information. An event is defined as *something that occurs* (LIRICS, 2007); *something that happens[10]; something that happens at a given place and time[11]; something that takes place[12]; anything that happens[13]*, etc.

Finally, the main drawback of definition (1) is the lack of linguistic encoding of participants (via arguments) and an event (via a verb). Strictly speaking, linguistic encoding cannot be omitted from the definition, as a semantic role is a linguistic concept related to other linguistic entities and, therefore, cannot be defined without their presence. As a result of hiding linguistic encoding, in some actual definitions of this type there is an entanglement of a) an argument and a participant; b) a verb and an event. Here are some examples with a highlighted mix of linguistic and non-linguistic notions[14]:

> *The semantic role is the role of **argument** in the **events*** (DLPP).
> *Semantic roles are essentially labels that are linked to arguments of verbs in order to identify the role each **argument** plays in the **event*** described by the verb (Valdimarsson, 2012).
> *A semantic role is the underlying relationship that a **participant** has with the main **verb** in a clause* (Payne, 1977).
> *Semantic roles are abstract models of the role an **argument** plays in the **event** described by the predicate* (Jurafsky and Martin, 2008).

Since this cluster of definitions is misleading as to what a semantic role should be, it will be excluded from any further consideration.

2.3. Semantic role applied to the syntactic level of a language

In this subchapter, we consider definitions of a semantic role on the syntactic level of a language. This means that in all cases the bearer of a semantic role is an argument (syntactic constituent).

There are two different types of definition on this level: a role as a relation and a role as an entailment. Two sub-clusters will therefore be analyzed.

---

[10] https://www.merriam-webster.com/dictionary/event
https://www.dictionary.com/browse/event
https://www.macmillandictionary.com/dictionary/british/event
[11] https://www.definitions.net/definition/event
[12] https://www.yourdictionary.com/event
https://www.thefreedictionary.com/event
[13] https://dictionary.cambridge.org/dictionary/english/event
https://www.collinsdictionary.com/dictionary/english/event
[14] In the definition of an event with linguistic encoding via a predicate and arguments, there is also an entanglement of arguments and participants: "An event includes a predicate (main relation, frame-evoking element), which is the main determinant of what the event is about. It also includes arguments (participants, core elements) and secondary relations (modifiers, non-core elements) (Abend and Rappoport, 2017).

2.3.1. Semantic role as a relation

This type of definition is the most popular and widespread. Since a semantic role is applied to an argument in a sentence, the process of role assignment is often called argument-indexing. There are three approaches to assign a semantic role to an argument depending on the nature of a relation.

2.3.1.1. Approach #1: a relation is unspecified

According to approach #1, a semantic role is considered as a relation of an argument to a verb in a sentence. What is important is that there is no indication to what the type of the relation is, i.e. the relation is unspecified.

Examples of actual definitions suitable for the cluster. In comparison with the other two approaches, this one includes the most numerous set of semantic role definitions:

> *A semantic role is a relation of arguments to the predicate in a clause* (Sugisaki, 2016).
>
> *A semantic role is the relation between a syntactic constituent and a predicate* (Ochoa et al., 2011).
>
> *A semantic role is the relation of an argument with a verb* (Perez, 2005).
>
> *A semantic role is the relationship that a syntactic argument has with the verb* (Lin et al., 2007).
>
> *A semantic role is the underlying relation that a constituent has with the main verb in a clause* (Mefteh et al., 2016).
>
> *A semantic role is the relationship that a syntactic constituent has with a predicate* (Shi et al., 2010), (Moreda et al., 2004).
>
> *A semantic role is the relationship that the syntactic argument has with the verb* (Priya and Kurian, 2011).
>
> *Theta-roles... label relations of arguments to predicators* (Rappaport and Levin, 1988:17).

Specific semantic roles extracted on the grounds of the definition. In 1968, Ch. Fillmore proposed the following roles: Agent, Patient, Instrument, Recipient, Result, and Location. Later, they were revised to eight roles: Agent, Experiencer, Instrument, Object, Source, Goal, Place, and Time (Fillmore, 1971). After that, the varieties of empirically extracted semantic roles ranged from a few to hundreds[15]. As a result, modern linguistic resources contain significantly different numbers of roles, which influence the process of semantic role annotation[16]. For example, VerbNet has 30 roles (v. 3.2) and 39 roles (v. 3.3), while FrameNet contains 656 roles (v.1.5) and 727 roles (v.1.7). Uncertainty in the role list led to semantic role labeling systems being created using Propbank or AMR resources[17], which define roles on a verb-by-verb basis without any further generalization.

Formal definition for the cluster. Despite the simplicity of the above-mentioned definitions, obtaining a formal definition requires preparation. First, the terms mentioned in definitions – *syntactic constituent, syntactic argument* or even simply *argument* – should be clarified. A syntactic argument is considered an argument in a particular syntactic position. In turn,

---

[15] See (Somers, 1987) and (Levin and Rappaport, 2005) for summary.

[16] See (Ellsworth et al., 2004) for details.

[17] Propbank (Palmer et al., 2005) uses 5 predicate-specific core roles and 13 non-core roles. The resource AMR uses the same number of core roles but 40 non-core roles (Banarescu et al., 2013).

syntactic position is the relative location of an argument with respect to a verb. It can be the subject, direct object, indirect object or object of a preposition. Arguments should be linked to their positions in a clause, otherwise there is a problem in role assignment. For example, in the event *hit (John, Peter)*, if the arguments *John* and *Peter* are not in positions of subject and direct object accordingly, there is no clue as to whether *John* is the Agent (*John hit Peter*) or the Patient (*Peter hit John*).

So, a syntactic argument can be denoted as a pair $< a_i, p_j >$, where:

--a_i is an argument;

--p_j is the position of a_i with respect to the verb;

Second, a syntactic argument occurs in a clause. For each clause, an argument a_i can only be in one position p_j and, vice versa, p_j contains just one argument, a_i. If there are n arguments in a clause, there are n pairs $< a_i, p_j >$. Each clause has a specific number of arguments and it varies in different clauses. The question is whether a clause should be an additional parameter in the representation of a syntactic argument: $a_i^c, p_j^c >$, where c is a clause. The answer is that it is redundant since the relation of an argument in a position towards a verb is not influenced by the other elements of a clause. If we know the argument, its position, and the verb, the clause itself does not matter in defining a semantic role.

Taking into account the above-mentioned preparations, a semantic role SR is a relation R a syntactic argument $< a_i, p_j >$ has with the verb V_k. The formal definition that reflects the actual definitions will look as follows:

(2) $$SR_{<a_i,p_j>}^{V_k} = R\left(< a_i, p_j >, V_k\right)$$

According to (2) a semantic role depends on three features: the lexical meaning of an argument a_i, its syntactic position p_j and the lexical meaning of a verb V_k. The following examples demonstrate how a semantic role can be changed if only one of these three features changes:

a) The argument is changed (*lemon* vs. *sister*). Position and the verb are the same:

John drinks tea with lemon.

John drinks tea with sister.

b) The position is changed (subject vs. direct object). The argument and the verb are the same:

John hit Peter.

Peter hit John.

c) The verb is changed (*hit* vs. *see*). The arguments and positions are the same:

John hit Peter.

John saw Peter.

Range of variables in formal definition. Syntactic arguments and a verb are variables presented in (2). Theoretically, there are four possible cases for the range of variables.

Case #1: a syntactic argument $< a_i, p_j >$ runs through all arguments (n_k) and a verb V_k runs through the set of all verbs $\{V\}$:

$$(< i, j >= \overrightarrow{1, n_k}) \ \& \ (V_k \in \{V\})$$

Case #2: a syntactic argument $< a_i, p_j >$ runs through some (m_k) but not all arguments and the verb V_k runs through the set of all verbs $\{V\}$:

$$(< i, j >= \overrightarrow{1, m_k}) \ \& \ (m_k \leq n_k) \ \& \ (V_k \in \{V\})$$

Case #3: a syntactic argument $< a_i, p_j >$ runs through all arguments (n_k) but a verb V_k runs through some set of verbs $\{V\}_R$ but not all of them:

$$(< i, j >= \overrightarrow{1, n_k}) \ \& \ (V_k \in \{V\}_R) \& \ (\{V\}_R \subset \{V\})$$

Case #4: a syntactic argument $< a_i, p_j >$ runs through some (m_k) but not all arguments and a verb V_k runs through some $\{V\}_R$ but not all verbs $\{V\}$:

$$(< i, j >= \overrightarrow{1, m_k}) \ \& \ (m_k \leq n_k) \ \& \ (V_k \in \{V\}_R) \ \& \ (\{V\}_R \subset \{V\})$$

The 4 above-mentioned cases can be summed up as follows:

	all arguments have roles	all verbs are eligible
case #1	yes	yes
case #2	no	yes
case #3	yes	no
case #4	no	no

Analysis. The prevailing approach in the theory of semantic roles is case #1, with a semantic role being applicable to any argument of any verb (Levin and Rappaport, 2005:36). Case #1 has been called the statement of *completeness* (Bresnan, 1982)[18]. There are two reasons behind case #1 and the follow-up statement of completeness. First, it is difficult to explain why some arguments do not have semantic roles. There is no obvious criterion as to why role assignment might happen in one case and not in another. Second, a semantic role delivered as an analog of a grammatical case on a higher level[19] was originally named a semantic case or deep case[20] and,

---

[18] There are the other statements that are more rigorous: coherence, distinctness, independence, uniqueness, Chomsky's bi-uniqueness, which create even more issues in argument role assignment.

[19] Classically, thematic roles are used in theories of case assignment ("grammatical linking") of verbs and perceived as something like the semantic counterpart of grammatical case… A direct theory will assume that cases of verb's arguments and thematic roles correspond to each other directly (Eckardt, 2013).

[20] To distinguish the Fillmorean cases from the usual case concept, they have sometimes been called *deep cases* (because Fillmore claimed that they were universally present at "deep structure"), *case roles*, or *case relations*, but it seems simpler and less confusing to call them *semantic roles*, a framework-neutral term that by now has wide currency (although it did not exist in the mid-1960s) (Haspelmath, 2006).

by default, should be applied to all arguments in the same manner as a grammatical case. As a direct consequence of the statement of completeness, semantic role theory is considered a full coverage theory. Meanwhile, there are multiple indications that straightforward extrapolation of syntactical cases to semantics via case #1 is questionable.

a) Range of verbs. There are at least four cases where verbs create an issue with role labeling. According to the first case there are many transitive verbs for which it is difficult to assign any easily characterizable, yet somewhat general, semantic roles from the most common semantic role inventories: *the engineer praised the bridge, the engineer touched the bridge, the engineer avoided the bridge, the engineer ignored the architect, the engineer praised the architect, the engineer greeted the architect,* etc. (Levin, 1999). The above-mentioned sentences force us to introduce verb-specific labels for participants such as *avoidee* or *praisee* and comparable *–ee* words. There are other verbs with the same predicament, like the verb *echo* in *The voice echoes through the forest* (Kittila et al., 2011) or the verb *promise* in *He promised money to the museum* (Lin, 2014). Newmeyer (2010:689) pointed out that in the sentence *I played a sonata*, the argument *sonata* is neither Theme nor Patient.

According to the second case, there are verbs with a perspective view on an event and corresponding perspective-dependent roles. For example, in *X buys Y from Z* the argument *X* is considered the Agent while in reality the event consists of two sub-events with two agents: *X buys* and *Z sells*. Sometimes the mapping of semantic roles to syntactic arguments is a matter of pragmatics or lexical stipulation: *he owns the car* vs. *the car belongs to him* (Davis and Koenig, 2000). Dowty pointed out that the sentences *The lamp is over the table* and *The table is under the lamp* or *X is to the left of Y* and *Y is to the right of X* describe exactly the same state of affairs. It means that such perspective-dependent notions as Figure/Ground and Gruber's stative Theme should be ruled out as candidates for thematic roles. (1991:563). In total, the case contradicts the idea of semantic roles as invariable entities that provide an unbiased (objective) view across various perspectives on the same event.

According to the third case:
(a) There are verbs that lexicalize arguments making them hidden. In this case, a semantic role is implicit and conflates with a verb. A verb can lexicalize a) path: *leave* vs. *move away from*; *rise* vs. *move upward*; *approach* vs. *come near*; *pierce* vs. *go into or through;* b) origin or goal: *unplug* vs. *move out of a socket*; *plug* vs. *move into a socket*; *pocket* vs. *put something into a pocket*; c) theme: *berry* vs. *gather berries*; *fish* vs. *catch fish; hay* vs. *cut grass*; d) theme and path: *paint* vs. *put paint on*; *butter* vs. *put the butter on*; *bottle* vs. *put liquid into bottle;* e) instrument: *sponge* vs. *wipe with a sponge*; *write* vs. *use a pen to mark; fly* vs. *use an airplane to move;* e) instrument and theme: *floss* vs. *clean between someone's teeth with floss;* etc. As a result, another participant (for example, *bread* in *John buttered the bread*) takes the role of lexicalized participant, creating confusion in role representation. Lastly, the lexicalization of arguments can lead to a situation when there is a verb with no arguments at all -- *it is raining* -- where the verb lexicalizes theme (*water*) and path (*downward*).
(b) There are delexical or light verbs, which are opposite to the previous type (a) since they have little semantic content of their own and use a follow-up noun to form a predicate: *have a*

nap, give a lecture, do a revision. Compare: *to give the pen* vs. *to give a hug*[21]; *to get the car* vs. *to get support*, etc. While the verbs from type (a) accumulate the surrounding arguments inside, the verbs from type (b) release their meaning on the surrounding arguments.

According to the fourth case, there are symmetrical predicates where both arguments seem to play the same semantic role. There is no motivatable role that can distinguish two arguments in: *John met with Peter, apples resemble pears, Italy borders France, A is similar to B, A is equal to B, A is near B, A weights as much as B,* etc (Dowty, 1991:556).

In the linguistic project VerbNet the issue with symmetrical verbs has been resolved by introducing the Co-role. For example, the verb classes *marry 36.2* (10 verbs) and *meet-36.3* (7 verbs) have the role frame [Agent+Co-Agent]; the verb classes *differ 23.4* (5 verbs) and *contiguous_location-47.8* (45 verbs) have the role frame [Theme+Co-Theme][22]. In the sentence *Italy borders France* Agent and Co-Agent are *Italy* and *France* while in the sentence *France borders Italy* Agent and Co-Agent are *France* and *Italy* accordingly. Such approach makes the criterion for choosing a semantic role purely syntactical.

Finally, there are verbs that govern other verbs. In the sentence *I like tomato* which is equivalent to *I like to eat tomato* the verb *like* has another verb *eat* as an argument. To accept this case, definition (2) should be extended to the verbs of the second order where a verb can be one of the arguments a_i with a semantic role assigned to it.

b) Range of arguments. An argument in a sentence can represent everything imaginable. It can be:

a) a physical object (*book* in *John was given a book).*

b) an abstract object (*lecture* in *John was given a lecture).*

c) a property (*redness* in *Skin redness frightened John).*

d) an event (*smoking* in *Smoking leads to cancer).*

As a result, the concept of a participant in an event becomes extremely broad. If the roles of physical participants or "roles of entities, such as peoples and things, involved in the situation" (Yule, 2006) don't raise any doubts, the roles of non-physical participants often lead to confusion. For example, if to accept that *joke* in the sentence *The joke made me laugh* is an agent then it is unclear how to assign roles in the sentence *John told the joke that made me laugh*.

The assignment of roles to arguments b)-d) contradicts with definition (2). It concerns the roles of time[23] (*5pm* in *John arrived at 5pm),* manner (*quickly* in **John quickly told** them), property *(red* in *Rose is red),* extent/distance/measure (*2 miles* in *Peter jogged for 2 miles*), amount (*three* in *I have three brothers*), frequency (*twice a week* in *Peter washed the car twice a week),* etc. The point is that the above-mentioned roles are determined not by a verb or

---

[21] Ditransitive light verbs create additional issue with the number of semantic roles. For example, the sentence *John hugged Mary* has 2-role structure while *John gave a hug to Mary* suggests 3-role structure.

[22] https://verbs.colorado.edu/verb-index/vn/class-h.php

[23] In (LIRICS, 2007) *time* is defined as "a participant that indicates an instant or an interval of time during which a state exists or an event took place", and *location* is defined as "a participant that represents the place where an event occurs, or a state that is true, or a thing that exists". Here both *time* and *location* are properties of the entire event.

argument position but strictly by the lexical meaning of an argument and remain unchangeable if the verb or position is replaced by the other[24].

The following test provides the validity of role assignment:

if an argument X fits the ISA template – *X is [role]* – it means there is a problem with role assignment.

For example, the argument *5pm* in the sentence *John arrived at 5pm* fits the template: *5pm is time;* the argument *2 miles* in the sentence *Peter jogged for 2 miles* fits the template *2 miles is a distance*, etc. In reality, *time* is not a role for *5pm,* it is a generic name for *5pm;* the same way as *distance* is not a role for *2 miles* but its generic name. Following this template for *Peter* in the sentence *Peter jogged for 2 miles* we should get *Peter is a person* but *a person* is not a role here. *Peter* is *agent*. A role is a horizontal (event-dependent) relation, not a vertical (hierarchical) ISA relation. Conceivably, the sentences with hierarchical relation like *A buffalo is an animal* are not related to role representation at all.

Being captivated by an obligation to assign a role to every argument, we have no choice but to produce rather strange and dubious semantic roles. A full range of arguments inevitably creates difficulties in role labeling[25].

2.3.1.2. Approach #2: a relation is specified as conceptual

The only difference of approach #2 from approach #1 is the clarification of the relation R as conceptual (semantic). Actually, the conceptual nature of the relation R was implicitly assumed in approach #1. Approach #2 just unfolds it. The other elements remain the same.

Examples of actual definitions suitable for the cluster.

> *Semantic roles are labels that identify arguments according to the semantic relation they bear to their verb* (Levin, 2005).
>
> *A semantic role is the conceptual relationship between the argument and the predicator in a clause* (Sanchez and Oliveira, 2008).
>
> *Semantic roles describe the relations that hold between a predicate and its arguments (e.g., "who" did "what" to "whom", "when", "where", and "how") abstracting over surface syntactic configurations* (Lang and Lapata, 2014).
>
> *A semantic role is the name of a semantic argument or the relation between a syntactic constituent and a predicate* (Weng et al., 2011).
>
> *Semantic roles—also known as thematic relations, theta roles, participant roles, and deep cases—are labels for certain recurring predicate-argument relations* (OBO, 2014).

---

[24] The same idea was stated in (Kittila and Zuniga, 2016): "The semantic link between verb and obliques is especially insignificant in the case of location and time expressions that are in principle common to all events regardless of their nature, because all events occur in time and space. This means that the semantic role of an oblique is determined by its own lexical semantics in these cases."

[25] There is another issue with arguments which is not related with their range. It is not clear which arguments are indispensable and which can be omitted. In different projects related to role systems there are core arguments and non-core arguments without any articulate criterion of their distinction. For example, in FrameNet the verbs *bend, break, cut, clear, remove, fill, freeze, open* do not have Instrument as a core element, but the verb *kill* has. On the contrary, in VerbNet *bend, break, cut, freeze, kill,* and *open* have Instrument in the role frame, but *clear, remove,* and *fill* don't have.

Specific semantic roles extracted on the grounds of the definition. The same as for approach #1.

Formal definition for the cluster. A semantic role SR is a conceptual relation R^C a syntactic argument $< a_i, p_j >$ has with the verb V_k :

(3) $$SR^{V_k}_{<a_i,p_j>} = R^C(< a_i, p_j >, V_k)$$

Range of variables in formal definition. Arguments and a verb as variables can vary according to the same four cases as in approach #1.

Analysis. The same as for approach #1.

2.3.1.3. Approach #3: a relation is specified through connection to the external world

Although approach #2 differs from approach #1 in the sense that it clarifies the nature of the relation as conceptual, both have something in common: no direct connection to the external world. Participants and an event are not presented explicitly on the right side of definitions (2) and (3). Approach #3 explicitly links together linguistic elements (arguments and a verb) and the elements from the external world (participants and an event).

Examples of actual definitions suitable for the cluster.

> *Semantic role is a function of a participant represented by NP, towards an event represented by a verb* (Huminski and Zhang, 2018).
> *Semantic role is a term used in syntax and semantics to refer to the semantic relations that link a predicate to its arguments in the description of a situation* (Mohammed and Ali, 2019).
> *Semantic role refers to the way in which the referent of the noun phrase contributes to the state, action or situation described by the sentence* (Finegan, 2007; Yule, 2006).
> *Semantic roles (also known as thematic relations or participant roles) are labels associated with the arguments of a verb by virtue of their part in the event denoted by the verb* (Levin, 2019).

Specific semantic roles extracted on the grounds of the definition. The same as for approach #1.

Formal definition for the cluster. A semantic role SR is a relation R a participant P represented by an argument $< a_i, p_j >$ has with an event E represented by a verb V_k:

(4) $$SR^{V_k}_{<a_i,p_j>} = R\left(P_{<a_i,p_j>}, E_{V_k}\right)$$

Range of variables in formal definition.

The variables on the right side of definition (4) -- participants and an event -- are in "signified-signifier" relation with arguments and a verb accordingly. So, the same four cases from approach #1 can be applied for approach #3. If, for example, the prevailing case #1 is accepted, then the variable $P_{<a_i,p_j>}$ runs through participants represented by all arguments; and the

variable E_{V_k} runs through events represented by all verbs. In other words, the variables of the definition (4) follow the same range as the variables of definitions (2) and (3)[26].

Analysis. In approaches #1 and #2, there is no need to define the variables – syntactic arguments and verbs -- since they are already well defined in linguistics. In approach #3, the variables have been changed and their definitions should be a prerequisite for the semantic role definition (4).

According to the lower indices for P and E in (4), the only way to do this is to define a participant as a signified of an argument and an event as a signified of a verb.

If events represented by all verbs are role-eligible and if participants represented by all arguments have roles, there is no need to care further about the definitions of a participant and an event. The statement of completeness allows linguists to avoid the answers to the following two questions:

> What is the nature of a participant to be role-bearing?
> What is the nature of an event to be role-eligible?

But in this case all the issues related to arguments and verbs (see 2.3.1.1.) will be inherited by participants and events. The notion of a participant will be extremely broad since an argument in a sentence can represent everything that can ever been imagined. A similarly bad situation occurs for an event since a verb as a signifier of an event lexicalizes arguments (which means an event sucks participants), can be delexical, represents perspective view, etc.

It is remarkable that in semantic role labeling (SRL), an event and participants are implicitly defined. Any SRL parser starts from the brief statement that the purpose of SRL is to find *who did what to whom, when and where*. It means that by default SRL deals with physical objects (*who, what, whom*) as participants and actions (*did*) as events.

2.3.1.4. Summary

Approach #1 operates with unspecified relation. Approach #2 specifies the relation by narrowing it to the conceptual one. Approach #3 specifies the relation by linking linguistic entities to the external world. In short, the existence of three approaches in the clustering of a semantic role as a relation can be illustrated by combinations of the two binary features: whether a relation is specified or not and if specified, whether it was done inside of a language or by linking to the external world.

	is relation specified?	if specified, it is done	
		inside of a language	by linking to external world
approach #1	no	N/A	N/A
approach #2	yes	yes	no
approach #3	yes	no	yes

---

[26] The project FrameNet (Fillmore et al., 2002) uses definition (4) with no restrictions on event diversity. All types of events are covered.

It is obvious that approach #1 with unspecified relation is too broad. Hereinafter, we will consider the definitions from approaches #2 and #3 only:

$$SR^{V_k}_{<a_i,p_j>} = R^C\big(< a_i, p_j >, V_k\big)$$

$$SR^{V_k}_{<a_i,p_j>} = R\left(P_{<a_i,p_j>}, E_{V_k}\right)$$

2.3.2. Semantic role as an entailment

According to the definition introduced by D. Dowty, a semantic role ("thematic role" in his notation) is considered not as a relation, but as a set of entailments. The main idea is that any complete semantic theory of natural language must permit to describe the *lexical entailments* of verbs. For example, if a sentence *x builds y* is true, then it is necessarily also true that *x* performs purposeful actions, that as a result of these actions a structure or other artifact *y* comes into existence, and so on (Dowty, 1989:75). A semantic role becomes a cluster concept that consists of a set of entailments.

Examples of actual definitions suitable for the cluster. There is just one definition of a semantic role related to the idea of entailment:

> *From the semantic point of view, the most general notion of thematic role (type) is a set of entailments of a group of predicates with respect to one of the arguments of each* (Dowty, 1991:552).

Specific semantic roles extracted on the grounds of the definition. Originally, five prototypically agent-like entailments and five patient-like entailments were empirically extracted[27].

Agent-like entailments include volitional involvement[28] in the event or state; sentience (and/or perception); causing an event or change of state in another participant; movement (relative to the position of another participant) and independent existence. Patient-like entailments include change of state; incremental theme; causally affected by another participant; stationary relative to movement of another participant and non-independent existence. These two sets of properties are called *Proto-agent* and *Proto-patient*, accordingly.

Clusters of certain properties correspond to the traditional roles and their variations. Agent is *volition + causation + sentience + movement*, or in some usages *volition + causation*, or just *volition* or, according to the ordinary language sense of Agent, *causation* alone. For example, Agent is not the same for the subjects of the verbs *murder* and *kill.* Property of volition is included for *murder* but excluded for *kill*. In general, Agent looks like a "floating entity" depending on the number of accumulated properties. Experiencer has only a property

---

[27] It was indicated that the properties are preliminary without implying that they are necessarily exhaustive or that they could not perhaps eventually be better partitioned in some other way (Dowty, 1991: 572).

[28] According to (Allen and Teng, 2018:236) volitionality is not signaled in linguistic usage, and should be derived based on commonsense knowledge and reasoning. For example, *John dropped the vase* can be understood in both ways: either intentional or unintentional.

of *sentience* without *volition* or *causation*. Instrument is *causation + movement* without *volition* or *sentience*. Theme is most typically represented as *change of state + incremental theme + non-independent existence + causally-affected*, but *causally-affected* is sometimes absent (Patient can be distinguished from broader Theme by this entailment) (Dowty, 1991:577).

Formal definition for the cluster. The formal definition requires some preparation.

Firstly, the notion of *entailment* is defined as a property (PR) of a syntactic argument[29] $< a_i, p_j >$ which takes place relative to the verb V_k:

$$PR_{V_k}(< a_i, p_j >)$$

Secondly, an *individual thematic role* (ISR) is defined as a set of entailed properties of a syntactic argument $< a_i, p_j >$:

$$ISR^{V_k}_{<a_i,p_j>} = \{PR_{V_k}(< a_i, p_j >)\}$$

Lastly, a third step is necessary to make a generalization of individual roles. A semantic role (SR) is defined as the intersection of all the individual thematic roles determined on the set of the verbs {V} with respect to one of the arguments $< a_i, p_j >$ of each:

$$SR^{\{V\}}_{<a_i,p_j>} = \bigcap_{\{V\}} ISR^{V_k}_{<a_i,p_j>} = \bigcap_{\{V\}} \{PR_{V_k}(< a_i, p_j >)\}$$

where the variable V_k runs through the set of verbs *{V}*.

Though not mentioned explicitly, it is obvious that entailments are made on the conceptual (semantic) level. Using a notation from (3), the previous definition can be transformed as follows:

(5) $$SR^{\{V\}}_{<a_i,p_j>} = \bigcap_{\{V\}} \{PR^C_{V_k}(< a_i, p_j >)\}$$

where $PR^C_{V_k}$ is a conceptual property of an argument which takes place relative to the verb V_k.

Range of variables in formal definition. Definition (5) raises the following 3 questions:

1. Does an argument $< a_i, p_j >$ run through the all arguments of the verb V_k ? The answer is negative. Proto-roles do not classify arguments exhaustively which means some arguments have neither role (Dowty, 1991:576).

---

[29] Dowty (1987:76) denotes an argument in position as a_i and uses the model-theoretical notation $V(a_1, a_2, …, a_n)$. We use the notation $< a_i, p_j >$ due to the following reason: ordering $(a_1, a_2, …, a_n)$ may not be the same as the linear order of the arguments in the clause. For example, for the verb *give* in the sentences (a) *John gave a book to Mary* and (b) *John gave Mary a book* the argument a_2 is uncertain. So, notation $V(a_1, a_2, …, a_n)$ assumes making some prearranged convention about the positions of the arguments.

2. Does a verb V_k run through all verbs? The answer is rather negative. There are verbs for which there may be no motivatable role to distinguish two arguments (Dowty, 1991:556).

3. How large should the set {V} be to define the intersection of properties? The question is open. Depending on the size of the set {V}, a different number of properties will be intersected. This number may either be too big to be useful or too small to be reasonable. For example, if the set {V} consists of just one verb, all properties related to the verb should be counted, which is obviously too specific. On the other hand, if {V} is the set of all verbs, the intersection should be empty.

Analysis. As shown in ch. 2.3.1, there are multiple indications that for the definition of a semantic role as a relation, the statement of completeness (case #1 as a prevailing one: all arguments have roles and all verbs are eligible for role assignment) is questionable. Negative answers on the above-mentioned questions 1. and 2. for the definition of a semantic role as an entailment contradicts the statement of completeness and confirm once again the opinion that not all arguments have roles and not all verbs are eligible for role assignment.

Presumably, because of the difficulties with answering question 3, Dowty didn't proceed further with definition (5) to define a set of semantic roles based on the idea of entailment. Instead, the gears have been changed to argument realization[30]. He developed the procedure to formally order the arguments as a subject and an object[31].

In actuality, Proto-agent and Proto-patient are not semantic roles. To say that an argument is Proto-agent is simply to state that by virtue of its possessing certain semantic properties, the argument will be selected as a subject in a sentence with that verb (Van Valin, 1999). For example, in the sentence *The stone rolled down the hill* the argument *the stone* has one property from Proto-agent (movement) while the argument *the hill* has one property from Proto-patient (stationary), so they are subject and object accordingly. But *the stone* is not an agent even if it has one of the properties of Proto-agent. It is a patient according to the traditional set of roles.

Nevertheless, proto-roles have an important consequence: a semantic role is not considered as a discrete entity but as a collection of entailed properties. According to definitions (2)--(4) there is an assumption taken by default that a semantic role is a discrete entity. Definition (5) states that the assumption is wrong and a role is not discrete.

Many publications appeared after Dowty's articles had been published concerning

---

[30] As Greg Carlson pointed out "a notion of a semantic role as a cluster of entailments of argument-positions of verbs was offered not as of mediating semantic interpretation but for selecting which argument for a verb should be placed in subject, and then object, position" (1998).

[31] Argument selection principle states that the argument for which the predicate entails the greatest number of Proto-agent properties will be lexicalized as the subject; the argument having the greatest number of Proto-patient entailments will be lexicalized as the direct object (Dowty, 1991:576).

critical comments[32], modifications of proto-roles[33], and practical usage of the idea of entailment[34].

2.4. Semantic role applied to the semantic level of a language

So far, semantic roles are defined as either a relation or a set of entailments. In both cases, semantic roles are tightly linked with syntactic arguments. Semantic roles are a kind of embedment of semantics into syntax, which bridges semantic and syntactic structures. This leads to a contradictory situation: without having in advance a semantic structure from which this embedment can be accomplished, semantic roles, as presumable elements of a semantic structure, are applied to the syntactic arguments. In other words, semantic roles are operated without knowing the semantic structure. As a result, the status of semantic roles remains vague. On one hand, they belong to the syntactic level since this is where the variables (syntactic arguments and a verb) of semantic role definitions are located. On the other hand, they are defined as conceptual relations (or conceptual properties) in definitions (3), (4) and (5).

One obvious solution to avoid the contradiction is to consider semantic roles as a direct semantic import onto a syntactically parsed sentence. Role labeling of a parsed sentence is treated as its meaning representation. In this case, there is no need to have a semantic structure in advance: a labeled sentence and its meaning representation are in 1-to-1 correspondence. As was shown in 2.3.1.1 and 2.3.2, this approach has multiple drawbacks and deviations.

To avoid drawbacks, another strategy for a semantic role has been suggested. The idea is that a semantic role is shifted completely to semantics and is considered an element of conceptual structure, not a part of syntax. Respectively, there is no parallel between conceptual structure and syntactic structure. They are far from identical.

We will consider two approaches or realizations of this strategy. One of them -- Schank's conceptual dependency system -- is more restricted and application-oriented while the other -- Jackendoff's conceptual semantics -- is more general and theoretical.

Schank's conceptual dependency theory (1975) is considered a model for natural language understanding. It is based on the idea that there should be a canonical meaning representation for natural language sentences. The representation is independent from the

---

[32] In general, there were 4 types of comments. (a) Critical opinions about validity and completeness of properties (Schlesinger, 1989:194), (Croft, 1998), (Levin and Rappaport, 2005), (Primus, 1999). For example, Croft pointed out that "there is no a priori reason why the Proto-agent and Proto-patient properties are partitioned in the way that they are, apart from inductive generalization from the empirical facts that Dowty adduces in support of his theory" (1998:37). (b) Properties of Proto-agent are not full-independent. Volitionality presupposes sentience. The property of existence is presupposed by the first 4 properties (Dowty 1991). (c) Properties of Proto-agent are not equal in determining argument selection (Dowty 1991:574). (d) Properties of Proto-patient are related to the properties of Proto-agent. See 3.2.2 for details.

[33] Properties of proto-roles have been modified and changed. For example, *incremental theme* has been replaced to *bounding entity* (Ackerman and Moore, 1998, 2001). Also *Proto-recipient* (Primus, 2006:55) and *Proto-goal/path/location* (Baker, 1997:120) have been added.

[34] In the project LIRICS for each role a list of associated entailments was made. Entailments were converted into a set of properties, e.g. [+/- intentionality], [+/- independent existence], etc. (Petukhova and Bunt, 2008:41).

words used in sentences, and for two or more identical (in meaning) sentences there should be only one meaning representation. Conceptual structures include 11 empirically identified primitive actions:

> PTRANS: to change the location of a physical object, e.g. *go*.
> ATRANS: to change the ownership, possession, or control of an object, e.g. *give*.
> MTRANS: to transfer information mentally**,** e.g. *tell.*
> MBUILD: to create or combine thoughts, e.g. *decide*.
> ATTEND: to direct a sense organ towards a stimulus, e.g. *listen, watch*.
> GRASP: to physically grasp an object, e.g. *clutch*.
> PROPEL: to apply a physical force to an object, e.g. *push*.
> MOVE: to move a body part, e.g. *punch, kick.*
> INGEST: to take something inside an animate object, e.g. *eat.*
> SPEAK: to produce a sound, including non-communicative sounds, e.g. *say*.
> EXPEL: to push something out of the body.

Schank states that eleven primitives[35] could represent most actions in the physical and mental worlds. In addition, it is assumed that each concrete action can be represented as a composition of primitives. For example, the sentence *John ate a frog* is represented as *John INGEST frog from unknown place to mouth by MOVE of John's hand from unknown place to mouth.*

Jackendoff's conceptual semantics (1983, 1987, 1990) is similar to Schank's dependency theory. The purpose of it is to extract the elements of conceptual structure by which a human understands language. Unlike Schank's 11 actions, Jackendoff introduced five primitive event-functions: GO, STAY, CAUSE, BE, ORIENT. Conceptual structure is considered a composition of these event-functions. For example, the same sentence *John ate a frog* is represented as a composition of CAUSE and GO: *John CAUSE a frog to GO into John's mouth*.

Examples of actual definitions suitable for the cluster. In both Schank's and Jackendoff's theories, empirically extracted primitives are the core elements of conceptual structure, while a semantic role is considered a derivative element.

According to Schank, semantic roles (*conceptual cases* in original notation) are dependents that are required by the ACT. A conceptual case is part of a primitive and is predicted by a primitive (1973a:198-205).

According to Jackendoff, a semantic role (*thematic relation* in original notation) is defined as *structural position in conceptual configurations* (1987:378). Later, it is identified as *an argument position in conceptual structure* (1993:31).

These two informal definitions from Schank and Jackendoff appear incomparable. Nevertheless, when transformed into a formal definition, both of them refer to the same concept and it is just a matter of terminology that makes them distinct.

Specific semantic roles extracted on the grounds of the definition. In Schank's conceptual dependency theory, there are five semantic roles: *Actor[36]*, *Objective*, *Recipient*, *Instrumental*

---

[35] The primitives are considered as independent, although it is not so obvious taking into consideration that, for example, INGEST is a particular case of PTRANS and GRASP assumes PROPEL.
[36] Actor is not explicitly considered a conceptual case.

and *Directive.* Each primitive requires either three or four semantic roles[37] (Schank, 1975:30). For example, PTRANS contains:

> *Actor* who initiates the PTRANS
> *Objective* which is PTRANSed (or moved)
> *Instrumental* which helps in moving;
> *Directive* which shows origin and destination of PTRANS.

The semantic roles associated with a primitive must be presented in conceptual structure, although some can be left empty (see formal definition for details).

In Jackendoff's conceptual semantics, the situation with semantic roles is different. There are no specific semantic roles since roles are considered *just convenient mnemonics for particularly prominent configurations* (1987:378). For example, Agent can be defined as the first position of the event-function CAUSE, Theme is the first position of the functions GO, STAY, BE and ORIENT; Source and Goal are the second (FROM) and third (TO) position of GO; Experiencer is presumably a position in a function related to mental states, etc.[38]. Also, the mnemonic character of semantic roles is rather convenient since for many argument configurations there is no traditional name. For example, in the sentence *John passes the house*, NP *the house* is not Theme, Goal, or Source (Jackendoff, 1990:46). Jackendoff's claim therefore is that semantic roles are not primitives in conceptual semantics.

Formal definition for the cluster. To set the formal definition of a semantic role on the conceptual level, one first needs to define the variables involved. Compared to the syntactic level with syntactic arguments and verbs, on the conceptual level there are variables of the other types.

Firstly, a verb V_k is replaced by a primitive $Prim_k$. Secondly, the conceptual structure is independent from the lexical meaning a_i of an argument $< a_i, p_j >$ since position p_j (or term *slot* used in conceptual dependency theory) can be left unfilled or empty. For example, in the sentence *John went to New York* with the following structure:
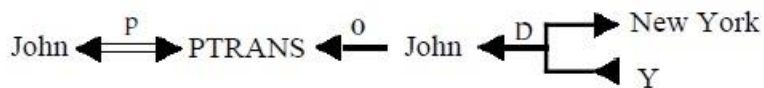

Figure 1. Example of Schank's conceptual structure

the position Y which represent Origin is not filled.

Regardless of the different terminologies and the different level of semantic role granularity (see 3.2. for details) used in Schank's and Jackendoff's theories, a semantic role SR can be defined as a conceptual relation R^C of position p_i to the primitive $Prim_j$:

---

[37] Actually, Schank mentioned "two or three semantic roles" but he didn't count *Actor*.

[38] As Jackendoff pointed out about roles, the constraints on their number and type follow from whatever constraints exist on the range of conceptual functions necessary to express the meanings of verbs (1987:379-380).

(6)
$$SR_{p_j}^{Prim_k} = R^C\left(p_j, Prim_k\right)$$

So, on the conceptual level there are two variables instead of the three used on the syntactic level.

Range of variables in formal definition. As for the range of positions, on the conceptual level everything is unfolded and there are no hidden positions. Every position is explicitly presented and has a semantic role which means that p_i runs through all positions of $Prim_j$.

As for the range of primitives, it is unclear whether all primitives are eligible for representation via semantic roles. Assuming that (a) a set of primitives covers all verbs, and (b) not all verbs are eligible for role representation, then it can be concluded that not all primitives are eligible for role representation. If a set of primitives covers only those verbs that are eligible for role representation, then there are no restrictions on the range of primitives in the definition (6).

So, theoretically, there are two possible cases for the range of variables:

Case #1: position p_j runs through all positions n_k and a primitive $Prim_k$ runs through the set of all primitives $\{Prim\}$:

$$\left(j = \overrightarrow{1, n_k}\right) \& (Prim_k \in \{Prim\})$$

Case #2: position p_j runs through all positions n_k but a primitive $Prim_k$ runs through some set of primitives $\{Prim\}_R$ but not all of them:

$$\left(j = \overrightarrow{1, n_k}\right) \& (Prim_k \in \{Prim\}_R) \& (\{Prim\}_R \subset \{Prim\})$$

Analysis. The replacement of verbs to primitives solves many issues related to verbs in definitions (3) -- (5) on the syntactic level. For primitives, there is no perspective view on an event, no argument lexicalization or, vice versa, delexicalization, etc.

Meanwhile, primitives have their own issues. The first is the lack of knowledge about the nature of primitives and the method of their extraction. While the set of verbs is finite and certain, we know very little about the set of primitives $\{Prim\}$, except for the trivial fact that the number of primitives should be much less than the number of verbs. As a result, the number of primitives on the conceptual level is uncertain. There are different sets of empirically extracted primitives and they vary from version to version. Schank (1973b) proposed 14 primitives and reduced them to 11 in the last version by eliminating SMELL, LISTENTO, LOOKAT and CONC, and adding ATTEND. Following the suggestions of Talmy (1985) and Culicover and Wilkins (1986), Jackendoff proposed the enrichment of five original primitives by introducing a new primitive function, ACT. Later, the function EXP for *X experiences Y* (Jackendoff, 1987:398) and the function AFF for *affect* with Actor in the first position and Patient in the second (Jackendoff, 1993:34) were added.

The second issue is related to primitive eligibility with respect to semantic roles: do all primitives include semantic roles as their components? As for the Schank's primitives concerns, the answer is positive: all 11 primitives are *action* primitives and they are eligible for role representation. States like *Peter is tall, Max is in Africa* or *The dog is on the left of the cat* are not represented through attributes with numerical scales. On the contrary, Jackendoff used primitive BE(x,y) to represent the above-mentioned states with corresponding roles, although the class of verbs covered by the primitive BE is questionable for role representation. See 2.3.1.1. (Range of verbs) for details.

What is the full list of primitives? Are all of them role-eligible? Both questions are unanswered.

2.5. Summary

The following table summarizes the most important elements of definitions (3), (4), (5) and (6):

level	definition	A semantic role is defined as a	Formal definition	Variables in formal definition and their range (variable:range)	
syntactic	1st definition	relation of argument and verb	$SR^{V_k}_{<a_i,p_j>} = R^C(<a_i, p_j>, V_k)$	verb:all	argument:all
		relation of participant and event	$SR^{V_k}_{<a_i,p_j>} = R\left(P_{<a_i,p_j>}, E_{V_k}\right)$	event, represented by verb:all	participant, represented by argument:all
	2nd definition	set of argument properties	$SR^{\{v\}}_{<a_i,p_j>} = \bigcap_{\{v\}}\{PR^C_{V_k}(<a_i,p_j>)\}$	verb:some (exact list is uncertain)	argument:some (exact list is uncertain)
conceptual	3rd definition	relation of position and primitive	$SR^{Prim_j}_{p_i} = R^C(p_i, Prim_j)$	primitive:uncertain	position:all

Table 1. Summary of semantic role definitions.

3. Formal definitions and the issues under investigation

The purpose of this chapter is to find out if the three formal definitions formulated in Chapter 2 can shed light on the two issues in semantic role theory formulated in Chapter 1 (Introduction):

Role inventory: how many semantic roles are needed and what they are.

Role combination: how to distinguish a valid set of roles from an invalid one.

3.1. Role inventory issue (1st issue)

The issue with semantic role inventory is directly related to the lack of knowledge about the role extraction procedure. As a consequence, the diversity of various role lists is purely empirical and inevitably leads to subjectivity and useless debates.

The issue with role inventory assumes, as a prerequisite, that roles are considered as primitives. Otherwise, the issue should be re-addressed towards entities that are considered as primitives. There is no sense in counting roles if they are compositions of "smaller" entities.

So, for each definition, two questions will be answered:

1) Is a semantic role considered a primitive in the definition?

2) If so, is there any procedure by which to extract semantic roles?

3.1.1. Definition #1

1) The definition of a semantic role as a conceptual relation (3) or by linking linguistic entities to the external world (4) is not transparent regarding the question of whether a semantic role is a primitive. The point is that the relation R is unspecified towards the level of verb generalization. It is unclear what set of verbs can specify the same relation R. For example, in the following two sentences -- *The wave hit Peter* and *John kicked Peter* -- at least three levels of verb generalization can be applied for the role assignment of *wave* and *John*.

First level (no generalization): *hitter* and *kicker* accordingly. For this level, the relation R is defined on a verb-by-verb basis without any higher generalization. The linguistic resource PropBank uses this level to define semantic roles (Palmer et al., 2005).

Second level (some generalization): *causer* and *agent* accordingly. For this level, the relation R is applied not to individual verbs but to a set of verbs and the specific lexical meaning of the argument is counted (*the wave* is not animate while *Peter* is).

Third level (higher generalization): *actor* for both subjects. For this level, the relation R is applied to an even larger set of verbs than in the second level since the specific lexical meaning of the argument is not counted (*the wave* and *Peter* are both actors regardless of whether one of them is animate while the other is not).

Taking into consideration the above-mentioned thoughts on verb generalization, we will further assume that definition #1 implies by default at least the second level of verb generalization and therefore a semantic role is considered as a primitive. Fig. 2 demonstrates the location of a semantic role in the transformation of the list of verbs into verb classes via role representation.

2) The definition doesn't give any clue on how to extract roles. An enormous variety of empirically extracted semantic roles indicates that there is no distinct criterion how to extract roles.
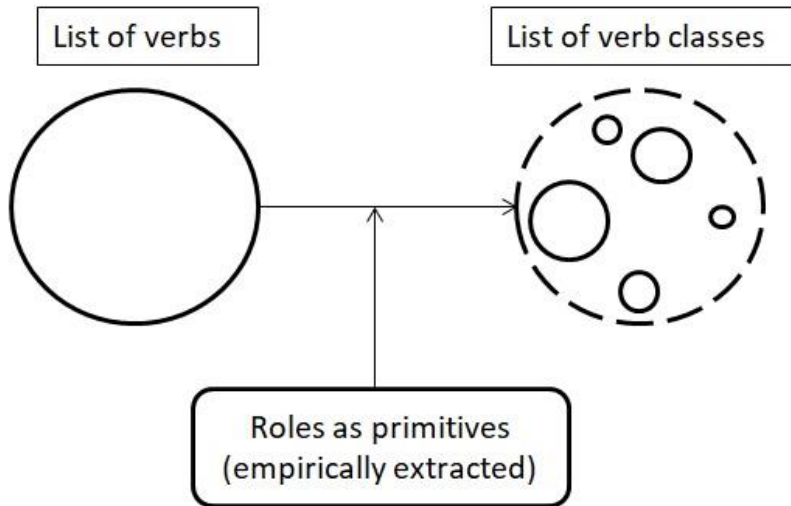
Figure 2. Roles as primitives in the definition #1.

3.1.2. Definition #2

1) Definition (5) points out that a semantic role is defined as a set of properties derived from entailments. It means that a role is definitely not a primitive and thus the issue with role inventory is denied. Instead, it is replaced with the no less difficult issue of property inventory: how many properties exist and what they are. It is obvious, definition (5) cannot help in solving this issue. Fig. 3 demonstrates the location of a semantic role and a property as a primitive in transformation of the list of verbs into verb classes.
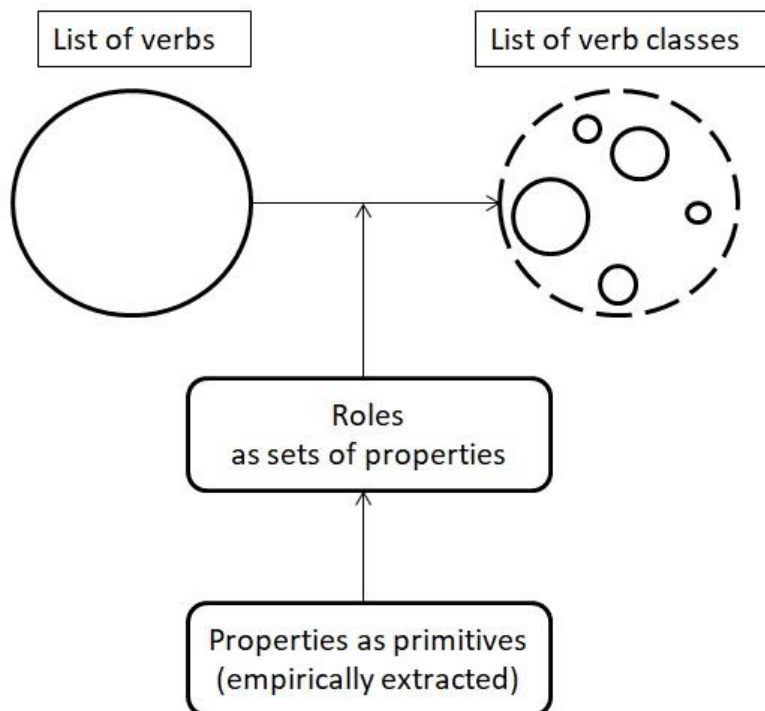
Figure 3. Properties as primitives in definition #2.

2) Not applicable since a role is not a primitive.

3.1.3. Definition #3

1) The status of a semantic role on the conceptual level appears more complicated. On the one hand, definition (6) already contains verb primitives, whether they are Schank's primitive actions or Jackendoff's event-functions. Verb primitives are extracted in advance and label certain classes of verbs. Therefore, the following can be stated:

(A) Verb primitives are primitives for some classes of verbs.

On the other hand, a semantic role cannot be defined as a composition of verb primitives in the same way a semantic role has been defined as a composition of entailed properties in definition #2. A composition of verb primitives is used for the representation of complex verbs. For example, as was mentioned in 2.4., the verb *eat* is represented as a composition of Schank's *INGEST & MOVE* or Jackendoff's *CAUSE & GO*. The phrase *some classes of verbs* in (A) means that verb primitives do not cover the full list of verbs. There are verbs which are supposed to be represented as a composition of verb primitives. Therefore, the following can be stated:

(B) Verb primitives are *not* primitives for semantic roles.

According to definition #3, a semantic role is defined as a relation between a primitive and a position. What is important is that a verb primitive can be represented via role combination in the same way the verb *hit* is represented via *hitter* and *thing hit* in Propbank. For example, Jackendoff's primitives can be represented as follows[39]:

GO == theme + origin/path/goal;

ACT == actor + actee;

CAUSE == causer + causee; etc.

In turn, Schank's primitives can be represented as follows[40]:

GRASP == grasper + graspee;

ATTEND == experiencer + stimulus;

ATRANS == theme + recipient, etc.

This means that, theoretically, the situation can be flipped: not a verb primitive defines roles but a combination of roles defines a verb primitive. In other words, if we know the list of primitives, we know the list of roles and vice versa. Verb primitives and their representation as a combination of roles are interchangeable. Therefore, the following can be stated:

(C) Verb primitives can be represented as a combination of roles.

Having statements (A) and (C) together, one can conclude that if a verb primitive is a primitive for verbs and if a verb primitive can be represented via roles, it means that roles are also primitives for verbs. Therefore, the following can be stated:

(D) Semantic roles are primitives for verbs in the same way as verb primitives.

---

[39] There is no proof here that all of Jackendoff's primitives are role-eligible (see 2.4. Analysis).
[40] Schank's original roles – *Actor, Objective*, *Recipient*, *Instrumental* and *Directive* -- are located on the higher level of generalization to count specifics of the primitive actions.

On the conceptual level, the list of roles does not exist in advance. It comes after the primitives, which are extracted empirically. Fig. 4 demonstrates the location of a semantic role in the process of verb transformation into a list of primitives.
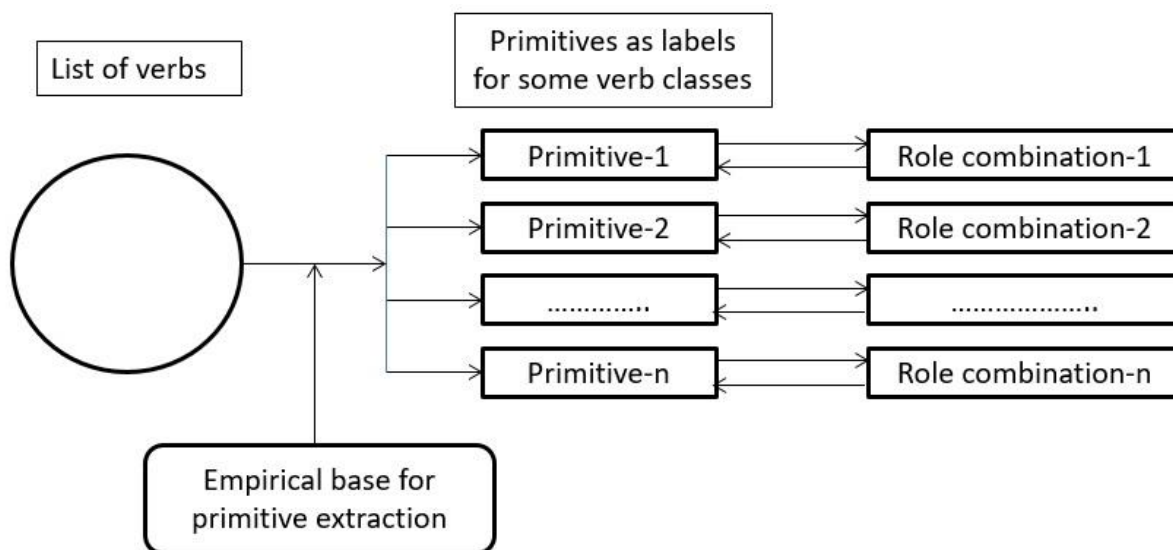

Figure 4. Roles as primitives in the definition #3.

2) In definition #3, the issue with role inventory is replaced by the issue with primitive inventory, since the extraction of roles depends on the extraction of verb primitives. Primitives in definition #3 face the same issue as roles in definition #1: a lack of criterion for extraction. Distinct role extraction from verb primitives has been neutralized by indistinct primitive extraction. As a consequence, a list of roles based on a list of verb primitives varies on the conceptual level.

3.2. Role combination issue (2nd issue)

The issue with role combination -- how to distinguish a valid set of roles from an invalid one -- assumes the presence of hidden relations between semantic roles. Three formal definitions will be examined to see if they can shed light on this.

It is important to underline in advance the difference between role combination and role hierarchy in the sense of commonalities[41] and differences across roles. For example, the role of Patient is definitely combined with the role of Agent, but according to role hierarchy, the Patient has much more in common with Goal than with Agent[42]. By analogy with syntagmatic

---

[41] There are many observations about roles commonalities. For example, Recipient is definitely a subtype of Goal; Goal and Benefactive share the same morphological case in some languages (Croft, 1991:157), etc.

[42] Riemer suggested the following thematic hierarchy that explicitly demonstrates the proximity of Patient and Goal: Agent > Beneficiary/Experiencer > Instrument > Theme/Patient >Goal/Source/Location. Agent will automatically be coded as subject. In the absence of Agent, any Beneficiary or Experiencer will be given subject status, and so on (2010:340).

vs. paradigmatic relations in linguistics, it can be said that role combinations are syntagmatic while the role hierarchy assumes paradigmatic relations[43].

The issue makes sense under the same obvious prerequisite as for the 1st issue: a semantic role must be a primitive.

3.2.1. Definition #1

According to the formal definition, a semantic role is defined as a relation between a syntactic argument and a verb. Fig. 5 illustrates this for the verb *break*. There is no evidence that semantic roles are somehow related to each other. It can be said they are unrelated.
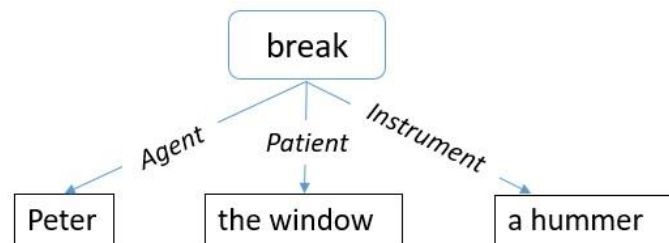

Figure 5. Semantic roles for the verb *break*.

As a consequence, roles in role frames for verb representation are flat (Lin, 2014:17), i.e. they are not aggregated in any specific order. For example, the frame [Agent, Instrument, Patient] in Fig. 5 is identical to the frame [Instrument, Patient, Agent]. It looks as if pieces of verb meaning are thrown into the same bag. The well-known term *role frame* is actually a disordered role list.

To check if roles are related, let us investigate empirically extracted roles. There are two observations that indicate the relation exists.

Firstly, if roles are unrelated, why do some role combinations refer to verbs while others do not? For example, if the role frame {Agent, Patient, Instrument} represents the verbs *break, cut, kill,* etc., or the role frame {Experiencer, Stimulus} represents *see, hear, taste,* etc., the frames {Patient, Experiencer}, {Location, Experiencer, Patient} or {Beneficiary, Goal} do not have any corresponding verb at all (Levin and Rappaport, 2005). In FrameNet (Fillmore et al., 2002) there are two types of role-to-role relations: *requires* and *excludes*, that indicate for a specific role in a frame which roles are required and which are excluded.

The discrepancy between valid and invalid numbers of role combinations are enormous. For example, a role frame for a verb in VerbNet (VN) contains a minimum of one role and a maximum of 6 roles. Having in total 30 roles in VN, the number of all possible role frame combinations equals 768211. In reality, there are only 107 valid role frames (Huminski et al., 2018). As Davis and Koenig point out, treating thematic roles as primitives that are unrelated to one another makes valid role combination purely accidental (2000).

---

[43] Deep cases are among the types of semantic relations that elements of sentence structures have with each other in context, rather than with the system of contrasts and oppositions that differentiate constituents paradigmatically (Fillmore, 1977:60).

Secondly, roles have their linked counterpart: Experiencer makes no sense without the presence of the cause or content of an experience (Stimulus) and Goal makes no sense outside of the kind of scene in which Theme moves toward some intended destination. (Fillmore and Kay, 1993). Faulhaber indicates that a semantic role is often dependent on another one, i.e. it cannot be regarded in isolation (2011:211).

There are exceptions, though. Not all semantic roles participate in a pairing relationship; for instance, Agent, which is often understood as a mutually dependent concept with Theme or Patient (Næss, 2007:37), can occur alone, generating a one-role frame (Perini, 2015:190). Agent alone occurs in sentences with intransitives:

John shaved (bathed, dressed)

where verb *shave* has the role frame [A]. Meanwhile, the previous sentence is identical (Dowty, 1989:94-95) to:

John shaved (bathed, dressed) himself

which is, in turn, structurally similar to the follow-up sentence with transitive:

John shaved (bathed, dressed) Peter

where the verb *shave* has the role frame with two roles of Agent (*John*) and Patient (*Peter*): [A+P].

There is another case with Agent alone on the syntactic level when the role of Theme is lexicalized by a verb. In the sentence *I berry* the only explicitly presented argument has the role of Agent since the verb *berry* lexicalizes the second argument *berry* as Theme[44].

The 1-role frame [Theme/Patient] is also allowable. Theme can occur alone, representing moving participant in sentences with the verbs *walk* or *run.* For example, *John walked the dog* and *John walked* are well syntactically*.* While it is amusing to say *John walked himself,* conceptually this is what we do mean (Schank and Riesbeck, 1981).

Patient alone occurs in sentences like *The rose bloomed* with the role frame [P]. If the role of Patient is assigned to a participant who is affected by an event, who or what affects *the rose*?

It is a common situation when certain arguments and corresponding roles are hidden (implicit) on the syntactic level[45], and because of that there are cases with isolated roles. In general, on the syntactical level, one cannot have fully-related roles.

In summation, the following conclusions can be drawn:

---

[44] In VerbNet there is verb class Berry-13.7 (29 verbs) with frame [Agent].

[45] An implicit argument is a conceptual argument that is neither expressed syntactically nor bound to an argument that is expressed syntactically. (Jackendoff, 1987:409)

1. There is no evidence from formal definition #1 that semantic roles are related.
2. Most of the empirically extracted roles show the presence of implicit relations between them.
3. There are unrelated roles (one-role frame), since some arguments and corresponding roles are hidden on the syntactic level.

3.2.2. Definition #2

The issue is not relevant to the definition since the entailed property of a syntactic argument, not a semantic role, is a primitive. However, the investigation of whether properties are related can shed light on the original issue with role combination.

The set of properties defines a semantic role, but there is no way to conclude from the definition if the properties are related. Therefore, in the same way definition #1, the empirically extracted properties that define Proto-agent and Proto-patient will be analyzed. As it was mentioned in 2.3.2., the list of five properties for each Proto-agent and Proto-patient were empirically entailed.

Firstly, by analogy with role combination from definition #1, one can conclude that not all combinations of Proto-agent and Proto-patient properties are valid. For example, the property *sentient* for Proto-agent does not match property *causally affected* for Proto-patient. There are some implicit restrictions on their combinations.

Secondly, although properties are hypothesized to be independently motivated (Dowty, 1991:572), there is no doubt that the properties for Proto-agent and Proto-patient are relational. The Proto-agent or Proto-patient status of an entailment of motion depends on the presence of another entity relative to which it moves. Similarly, causally affecting (a Proto-agent property) and being causally affected (a Proto-patient property) are correlated properties (Koenig and Davis, 2001:83). Croft indicates the complementary properties: volitionality–undergoer, causer–affected, moving–stationary (2012:191). A similar idea was expressed by Levin (2019:7-8), wherein the 3 Proto-agent properties are necessarily paired with Proto-patient properties.

Lastly, Proto-patient alone cannot be defined. If one participant of a predicate is causally affected, the predicate necessarily selects a causer as another participant. In other words, without Proto-agent there are no properties of Proto-patient (Primus, 1999:52)[46]. Dowty admitted that the next three entailments of Proto-patient -- causally affected; stationary relative to another participant; and existence not independent of an event --are the converses of Proto-agentive entailments ---- causation, movement; independent existence. If a verb has one of the first type of entailment for one argument, it necessarily has the corresponding one of the second type for another. (Dowty, 1991:574). In short, being solely Proto-patient does not

---

[46] For some reason, Primus states one-way dependency: Proto-patient entails the presence of Proto-agent, while Proto-agent does not. We guess it is 2-way dependency: a causee doesn't exist without a causer the same way a causer is pointless without a causee.

make any sense as its properties depend on and assume the presence of the properties of Proto-agent.

Dependency between Proto-patient and Proto-agent confirms the idea from definition #1 that one-role frames like [A], [P], or [T] occur on the syntactic level because some arguments and their corresponding roles are hidden.

In summation, definition #2 confirms the assumption the roles are related via the interdependence of empirically extracted primitive properties.

3.2.3. Definition #3

From the analysis of definitions #1 and #2, one would have expected that on the conceptual level the roles are fully related. Meanwhile, the situation is the same in formal definition #3 where there is no indication that the roles are related. However, as it was demonstrated in 3.1.3., empirically extracted primitives contain roles linked with each other.

Surprisingly, in the current state of art on the conceptual level, there exist some exceptions. It sounds strange, as it should not be the case, but there are original primitives with only one position. For example, the sentence *Bill rolled the ball down the hill* is represented by the 2-tier structure as follows*:*

GO ([BALL], [DOWN HILL])
ACT ([BILL], [BALL])

which is expected, but for the sentence *Bill rolled down the hill* (assuming that Bill did it volitionally) the original primitive ACT has only one position (Jackendoff, 1987:396):

GO ([BILL], [DOWN HILL])
ACT ([BILL]).

The sentence *Bill entered the room* has a similar structure (Jackendoff, 1993:34):

GO ([BILL], [TO[IN[ROOM]]])
AFF ([BILL])

where there is no Theme/Patient for the function AFF(affect) as if Bill affects nothing.
What should have been expected here is ACT ([BILL], [BILL]) and AFF ([BILL], [BILL]) accordingly (Bill as Agent ACT/AFF on Bill as Theme/Patient) since on the conceptual level the hidden roles are unfolded (see the analysis of examples for definition #1: *John shaved == John shaved himself; I berry == I gather berries; I walk == I walk myself;* etc.).

Therefore, the idea of role combination on the conceptual level is not consistent or fully realized. It can be explained by some traces or influence of syntactic structure; in particular, by the statement of *uniqueness* (every argument is assigned only one thematic role).

3.3. Summary

The following table uses the 3 definitions to summarize the key elements of the 1st and the 2nd issues:

level	definition	role inventory (1st issue)		role combination (2nd issue)	
		Is semantic role a primitive in formal definition?	Is there procedure to extract semantic roles?	Does formal definition indicate semantic roles are related?	Is there evidence that empirically extracted roles are related?
syntactic	definition #1	yes	no	no	yes
	definition #2	no	N/A	N/A	yes
conceptual	definition #3	yes	yes (from primitives)	no	yes

Table 2. Analysis of definitions regarding issues.

4. New Approach

In this chapter, a new approach to the semantic role is provided to answer the two issues under investigation.

4.1. Location of a semantic role

There is a discrepancy between a semantic role as an element of meaning representation and the historically dominant viewpoint of a semantic role as a semantic analog of a syntactic case.

As an element of meaning representation, a semantic role assumes that different but synonymous sentences have the same semantic role frame, while non-synonymous sentences have different role frames. Otherwise, the meaning representation is pointless.

As a semantic analog of a syntactic case, a semantic role inherits features of syntactic cases, the predominant one being completeness: a semantic role is applicable to any argument of any verb.

Completeness for syntactic arguments makes the concept of a participant represented by an argument extremely broad and leads to confusion in role assignment (see "2.3.1. Range of arguments" for details). It creates among Agent, Patient, Experience and Stimulus such dubious roles as Time, Attribute, Asset, etc. using another criterion and creating eclectic sets of roles[47].

Completeness for verbs[48] leads to a situation where role frame as meaning representation is either incomplete (for example, when a verb lexicalizes arguments and makes them hidden) or incorrect (due to a perspective view on an event or verb delexicalization), or simply unknown, as in the case of stative symmetric verbs (see "2.3.1. Range of verbs" for details).

Overall, the statement of completeness for a semantic role on the syntactical level forces us to consider a semantic role as a universal tool for meaning representation, a kind of "plaster for

---

[47] Due to necessity of completeness, Time, Attribute, Asset, etc. are presented in the projects Verbnet (Kipper Schuler, 2005), Propbank (Palmer et al., 2005), Framenet (Fillmore et al., 2002), or LIRICS (LIRICS 2007; Petukhova and Bunt, 2008:41).

[48] The ideal set of thematic roles should be able to concisely label the arguments of any relation (Bonial et al., 2011).

all sores", or the drug taken to be a panacea for all diseases. This assumption is too strong and wide to be real.

Analysis of definition #2 allows us to make the necessary step in comprehension that the range of verbs and arguments should be restricted, although the exact range is still uncertain.

This forces us to consider the 1st issue on the syntactic level as a wrongly formulated question[49] and to conclude that semantic roles are located on the conceptual level. It is easy to explain why it is useless to solve the issue with role inventory on the syntactic level: a transformation from the conceptual to the syntactic level means flattening conceptual structure into a linear set of arguments. It looks like a convolution of explicit and unfolded conceptual structure into something implicit and folded, which inevitably causes a distortion in role assignment. In the same way as one cannot discuss the form of saucepan handles if the entire saucepan has been flattened out, there is no 1-to-1 correspondence between conceptual structure and the structure of arguments on the syntactic level (see ch.6 "Appendix: comparison of semantic roles on the syntactic and conceptual level").

Completeness as a feature should not be applied to syntactic arguments and verbs that compromise themselves as variables for a semantic role definition, but instead to positions and verb primitives accordingly on the conceptual level[50].

The concepts of a position and a verb primitive will be analyzed in the following subchapters, but first, one needs to answer the obvious question regarding the formal definition on the conceptual level:

$$SR_{p_i}^{Prim_j} = R^C(p_i, Prim_j)$$

Comparing the definition with its syntactic counterpart

$$SR_{<a_i,p_j>}^{V_k} = R^C(<a_i, p_j>, V_k)$$

it is clear that a semantic role on the syntactic level depends on three elements: the lexical meaning of an argument, its position with respect to a verb, and the lexical meaning of a verb. Meanwhile, on the conceptual level it depends strictly on two elements: a position and lexical meaning of a verb primitive.

The question is: where is an argument of position and why does it not participate as a variable on the right side of the formal definition?

---

[49] "It has proved impossible to find a small set of roles that can be applied across all verbs in a language, let alone across languages. Yet this desideratum must be met if semantic roles are to figure effectively in accounts of linguistic phenomena" (Levin 2014). Although it is unclear from this statement what the procedure of the proof is, multiple attempts to do that and even the construction of a semantic role ISO standard (Bunt and Palmer, 2013) confirm the failure of the enterprise.

[50] This approach contradicts the definition of a semantic role in Oxford Bibliographies in Linguistics (Levin, 2014) which clearly indicates that it is about a verb and its arguments, not a verb primitive and its positions: "The meaning of a predicate, especially a verb, may be characterized via the relations that its arguments bear to it".

4.2. Argument on the conceptual level

On the syntactic level, pair $< a_i, p_j >$ indicates that a position itself is not enough to define a semantic role. A lexical item a_i is combined with its position p_j to build a relation with a verb for semantic role definition. On the conceptual level, an argument is not presented and it creates confusion. If there is no argument, what does a position alone mean?

In reality, on any level, whether syntactical or conceptual, a position must be filled. On the conceptual level, instead of arguments there is a category *THING* that underlies and substitutes any argument a_i in the pair $< a_i, p_j >$. Since there is no variations for *THING* in $< THING, p_j >$, the pair can be transformed as follows:

$$< THING, p_j > \implies < p_j > \implies p_j$$

where p_j is a position of the argument *THING*. Lexical item a_i that fits *THING* in a position p_j is located beyond the conceptual structure. The example of conceptual structure in Fig. 6 outlines the elements relevant to role definition (a primitive and two positions) on the conceptual level in the dash rectangle:
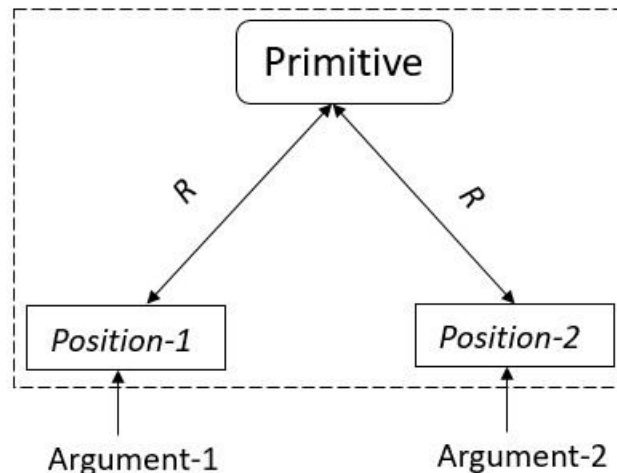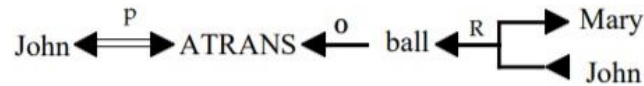

Figure 6. Elements (in dash rectangle) relevant for role definition on the conceptual level.

Traditional arguments are located outside of the dash rectangle. They replace the category *THING* if they fit a position[51] in the process of transformation from the syntactic level to the conceptual level. Some positions are filled by arguments, while others can be left without any replacement of the category *THING*. For example, the sentence *John went to New York,* according Schank's conceptual dependency theory (see Fig. 1. in 2.4.), has a conceptual structure with an open position of Origin since there is no corresponding argument for this position in the original sentence: *John went to NY from ___.*

---

[51] To fit a position an argument must have some inner properties that satisfy a position. For example, the position of *X* in *X perceives Y* must be filled by sentient argument. A stone lying on the road cannot be the case.

In general, the process of transformation is complicated. There are at least five cases for how an argument a_i can fill a position in the conceptual structure:

1. One argument fills one position: *John* and *Peter* in *John hit Peter* (see Fig. 6).

2. Several arguments fill one position: *John and Sam hit Peter*[52]*.* This case is illustrated in Fig. 7.

3. One argument fills several positions: *John* in the sentence *John hit himself*. This case is illustrated in Fig. 8. Filling two positions means that one argument has two semantic roles. For example, in the sentence *John threw the ball to Mary* represented as

$$\text{John} \xLeftrightarrow{\quad P \quad} \text{ATRANS} \xleftarrow{\; o \;} \text{ball} \xleftarrow{\; R \;} \begin{array}{l} \blacktriangleright \text{Mary} \\ \blacktriangleleft \text{John} \end{array}$$

in Schank's conceptual dependency theory, *John* is both Agent and Origin. There are many examples of combinations of Agent and Theme inside of one argument: *John* in the sentence *John went to New York*; *Mr. Wheeler* in *Mr. Wheeler jumped off the cliff* (Saeed, 1997:153); *John* in *John ran into the house* (Jackendoff, 1972) where *John* is simultaneously the initiator and the object of movement[53].
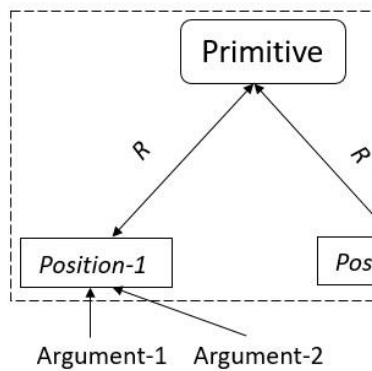

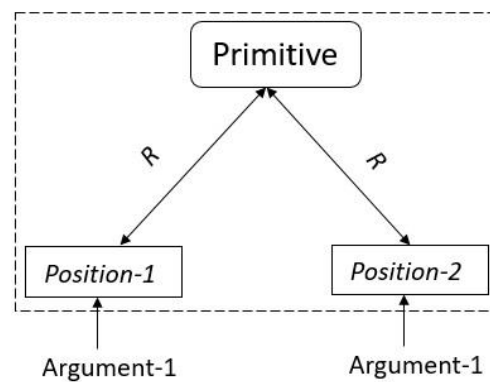
Figure 7. Several arguments fill one position.　　　Figure 8. One argument fills two positions.

4. An argument disappears on the conceptual structure: *hug* in *John gave me a hug*. The phrase *gave a hug* is used for a verb primitive. This case is illustrated in Fig. 9.

5. A new argument appears in the conceptual structure: *butter* in *Harry buttered the bread* (Jackendoff, 1987:387). This case is illustrated in Fig. 10.

As shown in Figs. 7-10, the rules formulated on the syntactic level regarding semantic roles are completely broken. Conceptual structure permits more than one argument for the same semantic role (Fig. 7); more than one semantic role for the same argument (Fig. 8); an argument without corresponding role (Fig. 9) and a role without an argument when it is lexicalized by a verb (Fig. 10).

---

[52] In the sentence *Five boys carried a piano up the stairs,* the phrase *five boys* represent Agent in the case of the group reading. In the case of the distributive reading, it is five events with one boy in each as Agent (Carlson 1998:9).

[53] The idea of two-role argument was elaborated by introduction of two tiers of roles: thematic tier and action tier (Jackendoff, 1990).

Regardless of the syntactic argument distribution (cases 1-5), the conceptual structure remains the same.
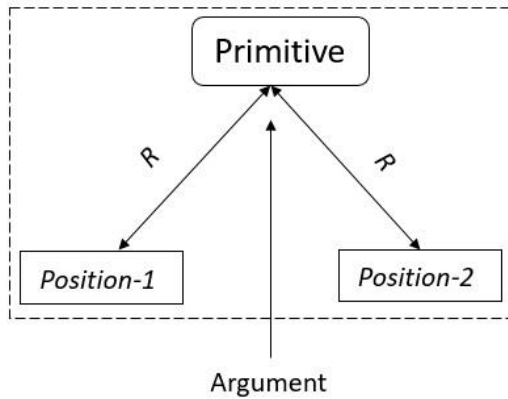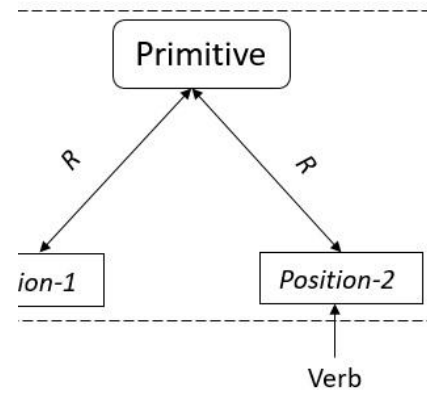

Figure 9. An argument disappears.

Figure 10. A new argument appears.

4.3. Position on the conceptual level

4.3.1. General view

Let us consider what position means on the conceptual level and whether it is possible to draw a parallel between positions on the syntactic and conceptual levels.

As follows from 4.2., all positions are filled by the category *THING*. During the process of transformation from a sentence to conceptual structure and depending on the argument variations in different sentences (see Fig. 11.), *THING* can be replaced or remain as it is in the conceptual structure.

The conceptual level should cover all possible argument variations in sentences with the same verb:
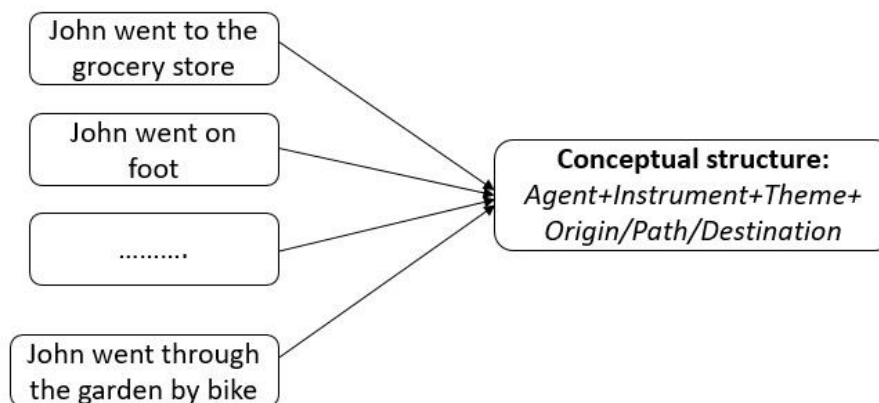

Figure 11. Transformation from a sentence to conceptual structure.

Otherwise, the transformation will not be possible.

As a result, positions on the conceptual level in comparison with those on the syntactic level are *mandatory elements.* They cannot be skipped or eliminated from the conceptual structure[54] in the same way as syntactical positions from the sentence.

4.3.2. The minimum number of positions

In previous chapters, we provided evidence that a role structure with one role does not exist on the conceptual level. It occurs because of the following reasons:

1. Empirically extracted roles from both the syntactical (definitions #1 and #2) and conceptual level (definition #3) are related (see Table 2 from 3.3.).
2. Exceptions with Agent, Theme or Patient alone (definition #1 in 3.2.) happen because some arguments and corresponding roles are hidden on the syntactic level. They are transformed into a 2-role structure on the conceptual level, since all roles must be presented there.

For example, in the project VerbNet there are four verb classes with 1-role structure:

VerbNet classes with 1-role structure	Role	number of verbs
berry 13.7	Agent	29
hiccup 40.1.1.	Agent	19
entity_specific_cos-45.5	Patient	41
weather-57	Theme	29

This is how a 1-role structure is transformed into a 2-role structure on the conceptual level:

- In the class *berry 13.7*, the verb lexicalizes Theme: *John berries = John picks berries*.
- In the class *weather-57*, the verb lexicalizes both Theme and Origin/Destination: *It's raining = water is falling from the cloud to the surface.*
- In the class *hiccup 40.1.1.*, if there is Agent, which entity does it influence? It is obvious that Agent and Patient are two inseparable elements inside of the human body: *John hiccups = Part of John influences on another part of John*.
- In the class *entity_specific_cos-45.5*, the situation is the same. In the sentence *The roses bloomed* with Patient (*roses*) alone, what entity initiates blooming?

---

[54] Whether or not certain syntactic cases are present in a sentence, the conceptual cases must always be there. That is, one could have the sentence: "Go." Conceptually, the ACT underlying "go" requires three conceptual cases (Directive, Objective and Instrumental) (Schank, 1975:30).

Finally, accepting 1-role structure, it is unclear how to distinguish an internal property of an argument (*The roses are red*) from its involvement in an event (*The roses bloomed*). Schank caught this difference and introduced the scales with numerical values (HEALTH, FEAR, MENTAL STATE, etc.) as a complementary approach to roles[55].

Let us consider the formal approach regarding the number of positions. If to assume that there is just one position, it becomes a constant, and formal definition (3) on the syntactical level is reduced as follows:

$$SR^{V_k}_{<a_i,p_j>} = R^C\left(< a_i, p_j >, V_k\right)$$

$$\downarrow$$

$$SR^{V_k}_{<a_i>} = R^C\left(< a_i >, V_k\right)$$

$$\downarrow$$

$$SR^{V_k}_{a_i} = R^C(a_i, V_k)$$

The reduction allows us to consider a semantic role as the relation between the lexical meaning of an argument and a verb. In the same way, formal definition (6) on the conceptual level is reduced as follows:

$$SR^{Prim_k}_{p_j} = R^C\left(p_j, Prim_k\right)$$

$$\downarrow$$

$$SR^{Prim_k} = R^C(Prim_k)$$

The last definition looks pointless since it considers a semantic role a unary relation of a primitive.

Therefore, one can conclude that a conceptual structure contains *a minimum of two positions*.

4.3.3. The difference between two positions

By analogy with the syntactic positions of subject and object for a verb, a conceptual position for a primitive:

- can be located before (similar to subject) a primitive;
- can be located after (similar to direct object) a primitive.

A position before primitive is considered a primitive starting point:

---

[55] See also the detailed analysis of role and scale representation in (Huminski et al., 2019).

For example, for the primitive *ACT*, a starting point is the answer to the question: *who/what ACT*. A position after a primitive is considered a primitive ending point:
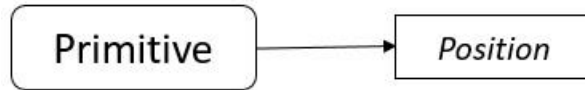
For the same primitive *ACT*, an ending point is the answer to the question: *ACT on whom/what*. A primitive does not have active or passive modes so the positions cannot interchange with each other. The entire picture looks as follows:
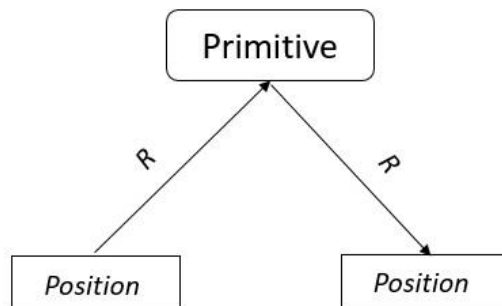
The only difference between two positions is their location (or direction/orientation) with respect to a primitive: it can be either before a primitive (direction: TO) or after a primitive (direction: FROM).

4.3.4. The maximum number of positions

If a conceptual structure contains a minimum of two positions, what is the maximum number of positions? According to 4.3.3., there is no way to define the third position since the definitions for two positions do not give a third possibility. Meanwhile, in Schank's conceptual dependency theory, each primitive requires either three or four semantic roles with a total number of five semantic roles: *Actor*, *Objective*, *Recipient*, *Instrumental* and *Directive*[56]. The cases with three and four positions will be considered.

4.3.4.1. Three positions

The appearance of the third position can be explained by a hidden causal relation between two primitives with two positions each. There are two cases here: a causal relation between either identical or different primitives. In both cases there are three positions -- X,Y,Z -- instead of two[57].

---

[56] It seems that in Jackendoff's conceptual theory each primitive has only two positions (Jackendoff, 1983:174; Jackendoff, 1987:375).

[57] Verbs are at most triadic (Levin Rappaport, 2005:44). This fact can be explained by assumption that verbs are represented via the causality of two primitives with two positions each.

A) Causal relation between identical primitives is represented as**:**

$$(X \, Prim_1 \, Y) \; CAUSE \; (Y \, Prim_1 \, Z)$$

where *Prim$_1$* is a primitive and the argument in the 2nd position of the left primitive (*Y*) is simultaneously the argument in the 1st position of the right primitive.

It allows us to represent two primitives with three positions that correspond to Agent (*X*), Patient (*Z*) and Instrument (*Y*). For example, the sentence *The doctor treated the patient by injection of antibiotics* represents this type of causation:

(doctor ACT antibiotics) CAUSE (antibiotics ACT patient)

Typically, there is a temporal pause between the action of the doctor and the action of the antibiotics, which is not permanently controlled by the doctor (Alexiadou and Schäfer, 2006).

On the contrary, the sentence *The doctor treated the patient with scalpel* represents a quick and tight causation where there is no spatial distance or time separation between the 1st and 2nd primitives. It is a kind of fusion between *Y* from the 1st and the 2nd primitives. Therefore:

(doctor ACT scalpel) CAUSE (scalpel ACT patient)

can be transformed into:

doctor ACT scalpel ACT patient

or, in general form, into:

X Primitive Y Primitive Z

By default, if nothing else in the place of *Y, X* can be there as *incorporated instrument* instead of Y:

X Primitive X Primitive Z

For example, in *John hit Peter*, John's hand serves as default (Jackendoff, 1987:401).

Causal representation of Instrument formally distinguish so-called instrument-causers *(antibiotics)* and pure instruments *(scalpel)* (Kamp and Rossdeutscher 1994)[58].

---

[58] There is a simple criterion to differentiate between a pure instrument and instrument-causer on the syntactic level. The former cannot be in the position of subject if Agent is omitted, while the latter can be:

> **The scalpel treated the patient.*
> *Chamomile treated the patient.*
> *The clock ticking woke the baby.*
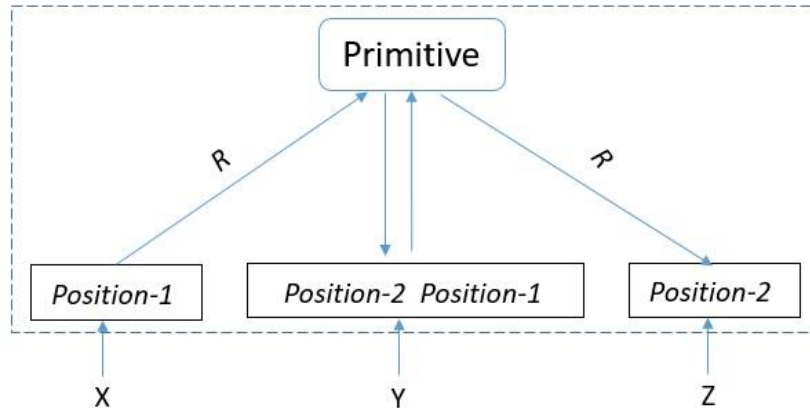> *The WW-2 mine wounded him* (Schlesinger, 1989).

Figure 12. Instrument as a fusion of two positions.

Fusion (see Fig. 12) represents Instrument not as a basic but as a *derived role* which is defined simultaneously as the 2nd (relative to *X*) and the 1st (relative to *Z*) position. That is to say, Instrument is a fusion of two roles: Patient and Agent[59]. It confirms the following three statements that assert that instrument is (a) intermediate, (b) inseparable and (c) derived:

(a) Instrument is a "participant in an event that is manipulated by an agent, and with which an intentional act is performed (Shiffrin and Bunt, 2007:38).

(b) Instrument is not defined separately since it is a manipulable force, i.e. incapable of independent motion and action and "under the control of another effector" (Van Valin and LaPolla, 1997:121);

(c) Instrument is not a new primitive role but an intermediary between Actor and Patient in the decomposition of an action (Jackendoff, 1987:401).

The formal definition of Instrument as a derived role extends the classical notion of instrument as applied only to action (primitive *ACT*) with Agent and Patient involved. Following the formal representation of causal relation between two identical primitives, Instrument can be applied to other primitives, as well. For example, there is no physical action in using *telescope* to increase the vision of the naked eye or *hearing aid* (or *earphone*) to improve hearing.

B) Causal relation between different primitives is represented as:

$$(X \; Prim_1 \; Y) \; CAUSE \; (Y \; Prim_2 \; Z)$$

where *Prim$_1$* and *Prim$_2$* are different primitives.

Three positions correspond to Agent (*X*), Patient (*Y*) and Directive/Origin/Destination (*Z*). For example, the sentence *John tossed the ball to the wall* represents this type of causation:

---

[59] That is why sentences like *The key opened the lock* are not freely usable in English. Speakers presented with this sentence in isolation generally have a clear intuition that *the key* is being given some contrastive force (DeLancey, 1991).

$$(\text{John ACT ball}) \text{ CAUSE } (\text{ball MOVE wall})$$

with *wall* as Destination.

In summation, a primitive with three positions is actually a causation of two primitives with two positions each.

4.3.4.2. Primitives with four positions

Four positions/roles occur in the case of hidden causation between three primitives with two positions each:

$$(X \text{ Prim}_1 \text{ } Y) \text{ CAUSE } (Y \text{ Prim}_1 \text{ } Z) \text{ CAUSE } (Z \text{ Prim}_2 \text{ } W)$$

The sentence *John picked up the crate with the crane to the 2nd floor* represents this type of primitive composition:

$$(\text{John ACT crane}) \text{ CAUSE } (\text{crane ACT crate}) \text{ CAUSE } (\text{crate MOVE } 2^{nd} \text{ floor})$$

It is common that *Y* coincides with *X* (incorporated instrument):

$$(X \text{ Prim}_1 \text{ } X) \text{ CAUSE } (X \text{ Prim}_1 \text{ } Z) \text{ CAUSE } (Z \text{ Prim}_2 \text{ } W)$$

The sentence *John (X) with his hand (X) tossed the ball (Z) into the wall (W)* represents this particular case.

4.3.4.3. Schank's five positions (roles)

Schank's five roles can be summarized as follows:

1. A primitive with three roles -- *Actor*, *Objective* and *Instrumental* -- can be represented via the causation of two identical primitives;
2. A primitive with three roles -- *Actor*, *Objective* and *Directive* – can be represented via the causation of two different primitives;
3. Recipient, as the receiver of an object, is a variation of Goal. The only difference between Recipient and Goal is that Recipient is animate, which is an internal property, not a relational one.

4.3.5. Conclusion

Taking into consideration that a primitive contains a minimum of two positions (4.3.2.) and a maximum of two positions (4.3.4.), one can conclude that the number of positions for a primitive is equal to two and only two.

4.4. Primitive on the conceptual and syntactic level

In comparison to the list of verbs, the list of verb primitives is uncertain. Their extraction procedure is unknown and, taking into account that some verbs cannot be represented through roles, not all empirically extracted primitives can be represented through semantic roles. There are well known doubts that the nature of some of them require another paradigm of representation (see 2.3.1.1. Range of verbs). In this chapter, we try to find the answer to the following question: is there any criterion to validate a primitive as role-represented?

4.4.1. Primitive on the conceptual level

As follows from the conclusion in 4.3.5., to be role represented a primitive must have two positions. This is the criterion for a primitive on the conceptual level.

4.4.2. Primitive on the syntactic level

The analysis of an argument on the conceptual level (4.2.) shows that it is not relevant in role definition. An argument a_i can fill the position in different ways without any influence on the meaning of a primitive and the conceptual structure itself.

Now, in a situation with a primitive and two positions, let us consider how positions can be filled by arguments.

Overall, there are two basic cases:

1. The 1st and the 2nd positions are filled by the same argument X, as shown in Fig. 13.
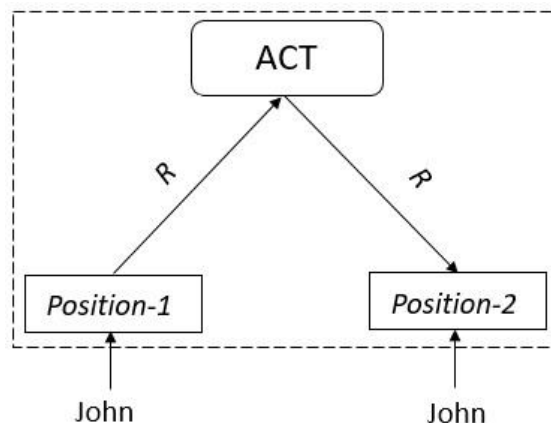

Figure 13. One argument fills two positions.

One argument filling two positions is considered a binary reflective relation. Optionally, it can also be a unary relation. For example, in the case of a binary reflective relation, the primitive *ACT* can represent the sentence *John hit himself*, assuming that one inseparable part of John hits another inseparable part (see the top right picture in Fig. 14). In the case of an optional unary relation, the primitive *ACT* can represent the sentence *John hiccupped/blushed/yawned*,

assuming that it is impossible to extract inseparable parts of John acting on each other (see the bottom right picture in Fig. 14).
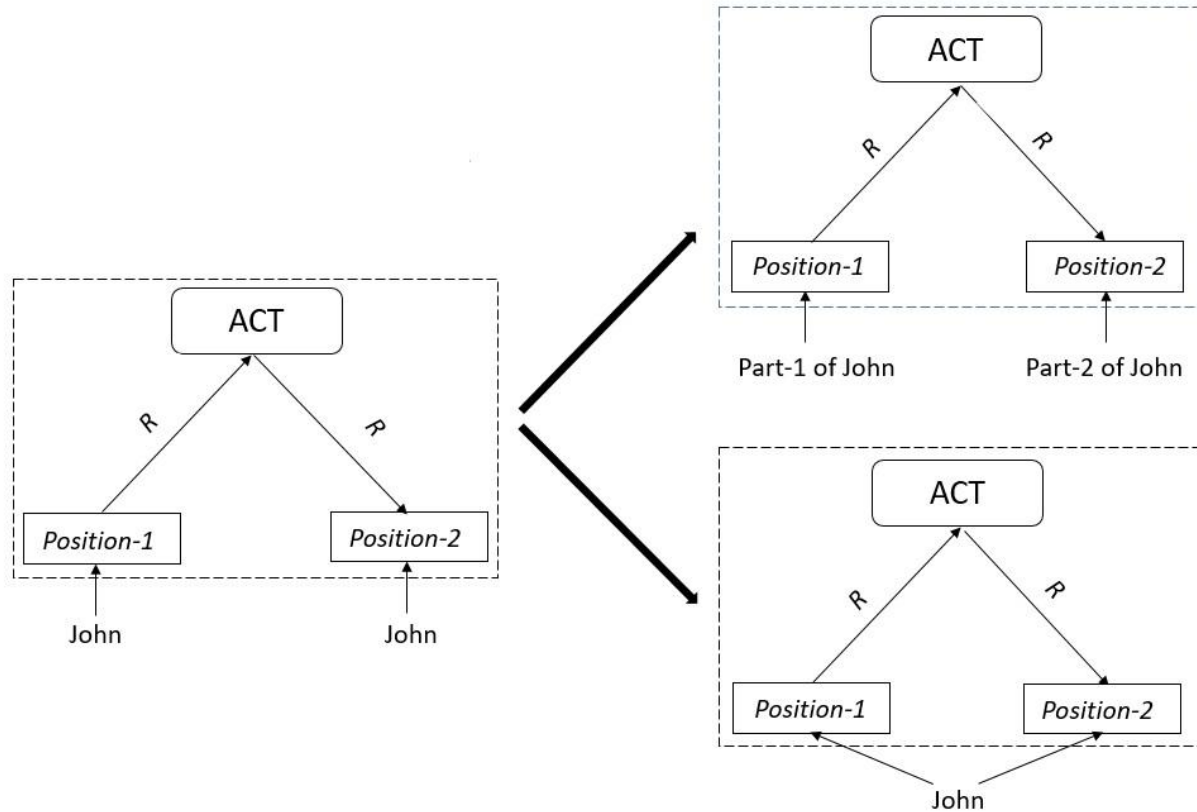

Figure 14. Variations for "one argument-two positions".

2. The 1st and the 2nd positions are filled by different arguments X and Y accordingly, forming a binary relation. The same primitive *ACT* can represent the sentence *John hit Peter* as binary relation (Fig. 15).
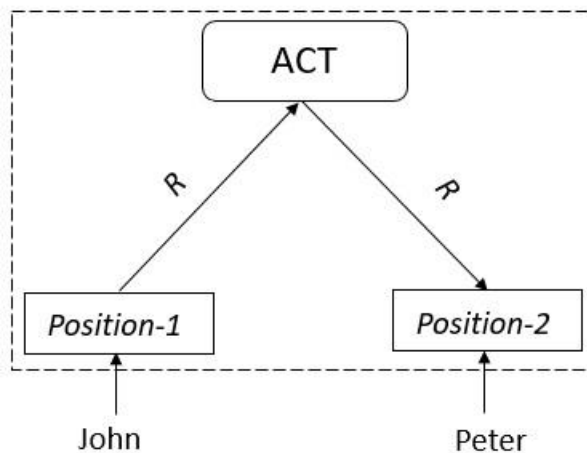

Figure 15. Two different arguments fill two positions.

There remains one unanswered question for this case. As was shown in the 1st case, one argument can fill two positions at once, which means that the binary relation in Fig. 15 shows uncertainty regarding the question of whether a reverse order of arguments is possible (see the left side of Fig 16).
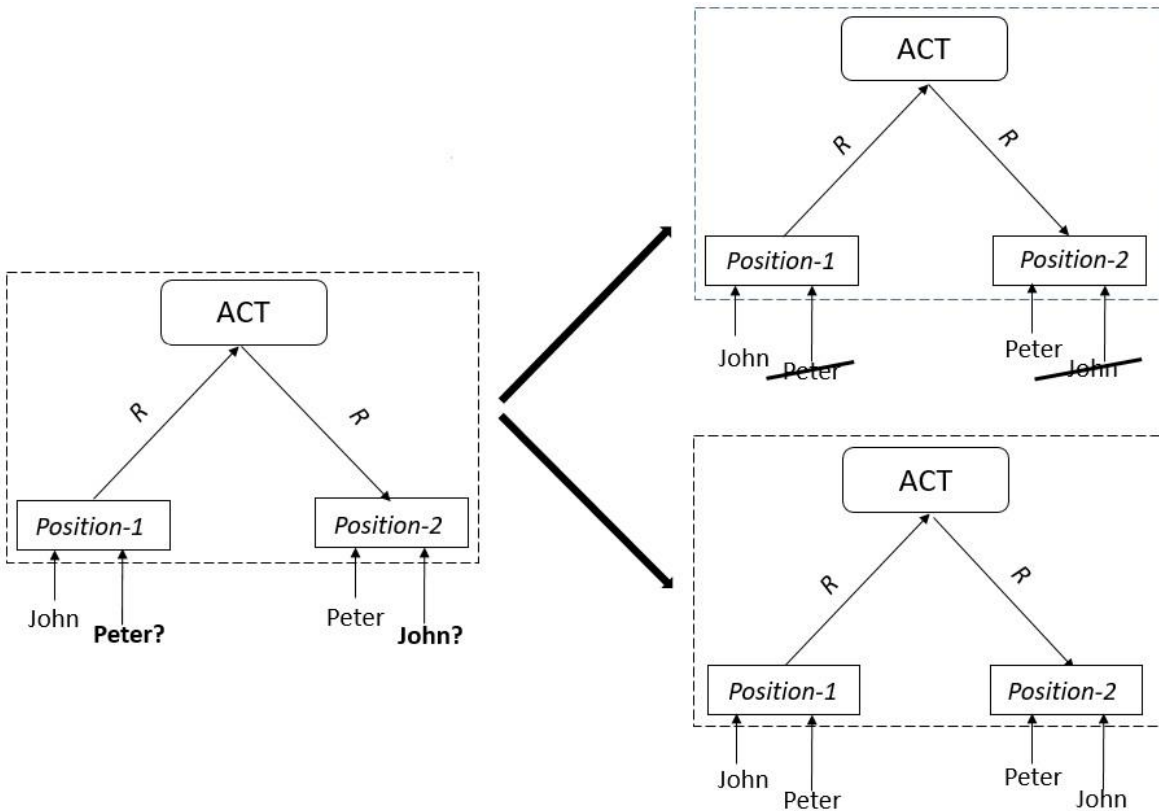

Figure 16. Two different arguments fill two positions with uncertainty of reverse order.

Two outputs are possible:

2a) the reverse order of arguments is not valid which results in a *binary asymmetrical relation.* In Fig. 16, the picture on the top right represents the sentence *John hit Peter but Peter didn't hit John*.

2b) the reverse order of arguments is valid and results in a *binary symmetrical relation*. This means that the structure with positions filled by two arguments – *John* and *Peter* – is complemented by the same conceptual structure with the same arguments that are flipped -- *Peter* and *John*. In Fig. 16, the picture on the bottom right represents the sentence *John hit Peter and Peter hit John == John and Peter hit each other.*

What is important is that there is no contradiction in having both 2a) and 2b) combinations:

(John hit Peter) & ¬ (Peter hit John);
(John hit Peter) & (Peter hit John).

where ¬ is a sign of negation. Of course, if the argument *Y* cannot fill two positions due to its lexical meaning (for example, *the ball* in *John hit the ball),* the combination with reverse order is not possible, but this is not the case. The case is in an opportunity for an argument to fill two positions if its lexical meaning allows for that[60].

To sum up, **the criterion** for a two-position primitive to be role represented is its property on the syntactic level to be reflexive, asymmetrical and symmetrical at once and unary (optional):

> *Prim(x,x)*
> *Prim(x,y) & ¬ Prim(y,x)*
> *Prim(x,y) & Prim(y,x)*
> *Prim(x)*

where *Prim* is a primitive, and *x* and *y* are arguments. For example, the primitive *perceive* on the syntactic level can be reflexive: *John perceives himself (John looks at himself)*; asymmetrical: *John perceives Peter (John looks at Peter but Peter does not look at John)*; symmetrical: *John perceives Peter and Peter perceives John (John and Peter look at each other)*.

The criterion formally explains why the following predicates are not role-represented:

1. *The lamp is over the table; the table is under the lamp; X is to the left of Y; Y is to the right of X,* etc.
2. *Bill married Kathy, apples resemble pears, X differs from Y, X borders Y, X is similar to Y, X is equal to Y, X is near Y, X weights as much as Y,* etc.

They are either asymmetrical (*is over, is under, is to the left of, is to the right of, is shorter than,* etc.) or symmetrical (*marry, resemble, differ, border, is similar to, is equal to, is near,* etc.). From this perspective, Jackendoff's primitives BE (e.g., *Carl is in the pub*), STAY (*Bill stayed in the kitchen)*, or CAUSE (*The crash caused hundreds of victims)* cannot be represented via roles for to the same reason.

4.5. Extraction of primitives

From 4.4., we conclude that to be role represented, a primitive must:

- a. have two positions on the conceptual level;
- b. be a binary relation which is simultaneously reflexive, asymmetrical and symmetrical on the syntactic level. It can be a unary relation, as well.

Let us call statements a. & b. *the criteria* for a primitive to be role represented.

In this chapter, *the criteria* will be used in the procedure of primitive extraction.

4.5.1. The procedure of extraction

---

[60] Carlson argues that the verb *kick* is asymmetrical since if John kicks the mule, it does not follow that the mule also kicks John (1998). Following the same logic, one can argue that the verb *kick* is symmetrical since if John kicks Peter it can follow that Peter also kicks John.

The procedure is based on the assumption partly presented in 4.3.4. that each primitive has at least one causal relation with another primitive. The causal relation between different primitives looks as follows:

(7) $(X\ Prim_1\ Y)\ CAUSE\ (Y\ Prim_2\ Z)$

where X, Y, Z are positions and $Prim_1$ and $Prim_2$ are two different primitives.
The idea is that by having one explicit primitive, another primitive can be restored. Namely:
(I)   if we know $Prim_1$, unknown $Prim_2$ can be restored:
   1. as a causee of $Prim_1$ by placing the argument of the 2nd position of $Prim_1$ (i.e. *Y*) in the 1st position of unknown $Prim_2$;
   2. by checking the criteria for $Prim_2$ to be role represented.
(II)   if we know $Prim_2$, unknown $Prim_1$ can be restored:
   1. as a causer of $Prim_2$ by placing the argument of the 1st position of $Prim_2$ (i.e. *Y*) in the 2nd position of unknown $Prim_1$;
   2. by checking the criteria for $Prim_1$ to be role represented.

4.5.2. Extracted primitives

To start extraction, we need to have at least one explicit primitive that does not raise any doubts to be role represented.

The primitive *ACT (physical)*

There is a well-known primitive that represents the class of verbs related to physical action: *hit, knock, kick, push, pull, press, shove,* etc. That is why it will be called *ACT (physical).*
Checking the criteria:
   a) On the conceptual level, there are two roles for the primitive that correspond to two positions: *who (what) act* or *Actor (physical)* and *act on who (what)* or *Actee (physical).*
   b) On the syntactic level, the primitive can be unary or binary reflexive, asymmetrical and symmetrical: *X ACT; X ACT on X; (X ACT on Y) &¬ (Y ACT on X);* and *(X ACT on Y) & (Y ACT on X)*. See examples for each case in 4.4.2.

The primitive *MOVE*

Assuming that *ACT (physical)* is $Prim_1$ in (7), what $Prim_2$ does it entail? It is obvious that *ACT (physical)* causes either movement or change. The argument in the role of *Actee (physical)* represents in $Prim_2$ a participant to be moved or to be changed. Therefore, the primitive $Prim_2$ can be called *MOVE (+change)*. It represents the class of verbs related to physical movement or change.
Checking the criteria:
   a) On the conceptual level, there are two semantic roles for the primitive *MOVE* that correspond to two positions: *what moves* or *Moving object* and *moves relative to what*

or *Object relative to which movement occurs*. The 2nd position of *MOVE* represents a semantic role that refers to the three traditional roles: *ORIGIN, PATH, and DESTINATION*. All three are particular cases of the general role *Object relative to which movement occurs*.

b) On the syntactic level, the primitive can be:

➢ asymmetrical and symmetrical: if *X MOVE relative to Y* it is possible to have both asymmetrical and symmetrical cases: *Y not MOVE relative to X* and *Y MOVE relative to X*.

➢ reflexive: *X MOVE X.* If the same argument fills both positions, we will get movement with respect to the same object, which is represented by the verbs like *rotate, raise (hand), bend (body),* etc.

➢ unary relation: *MOVE (x)* represents all types of changes of *X*. It is movement on the micro-level when it is impossible to extract the parts that move with respect to each other. There is no more specification here[61].

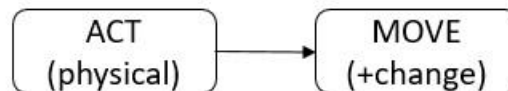We have the following causal relation (see Fig. 17):

Figure 17. Causal relation between ACT (physical) and MOVE (+change).

Example: the sentence *John threw the ball to the wall* is represented as *(John ACT (physical) ball) CAUSE (ball MOVE wall),* where *John* is *Actor*, *ball* is *Actee* and *Moving object* simultaneously, *wall* is *Object relative to which movement occurs* (in the traditional view it is Destination).

The primitives *ACT (cognitive)* and *ACT (emotive)*

There is no primitive after *MOVE* to be entailed since *Object relative to which movement occurs* is the role of a physical object standing still. So, we look at the *ACT (Physical)* to consider what primitive(s) can cause it[62].

Obviously, if *ACT (physical)* is performed by natural forces such as a hurricane, blizzard, rain, earthquake, tsunami, mudslide, volcano, flood, river, stream, tide, falling brick, etc., there will not be anything before the action. In the sentence *Avalanche hit the people*, *avalanche* is *Actor* and there is nothing behind it since the forces of nature are self-sufficient.

---

[61] Grimshaw pointed out that the difference between *melt* and *freeze* is probably irrelevant from the perspective of linguistics and that both verbs probably assign the same thematic roles (1993). See also (Huminski et al., 2019) about scale representation for verbs.

[62] The same primitive *ACT (Physical)* which causes the original *ACT (Physical)* indicates Instrument (see 4.3.4.1.).

The primitive *ACT (physical)* is what inanimate objects have in common with animate objects. The difference between them is that only animate objects can have something behind *ACT (physical)*. The following two primitives that can make animate objects to implement *ACT (physical)*:

-- a primitive *ACT (cognitive)[63]* that represents the class of verbs related to cognitive process: *think, understand, know*, etc.

-- a primitive *ACT (emotive)* that represents the class of emotional verbs: *enjoy, admire, like, love, fear,* etc.

Checking the criteria:

a) On the conceptual level, there are two semantic roles for the primitive *ACT (cognitive)* that correspond to two positions: *who act* or *Actor (cognitive)* and *act on who* or *Actee (cognitive)*. In the same way as for *ACT (emotive),* two semantic roles correspond to two positions: *who act* or *Actor (emotive)* and *act on who* or *Actee (emotive)*.

b) On the syntactic level, the primitive can be:

➢ asymmetrical and symmetrical. If *X thinks about Y* it is possible to have both *Y doesn't think about X* and *Y thinks about X*. The same way if *X hates Y* it is possible both *Y does not hate X* and *Y hates X*.

➢ reflexive: *X thinks about himself* and *X hates himself.*

➢ unary relation (optional): unknown.

Both of these primitives (see Fig. 18) make *Actor* an *intentional or volitional Actor*, which reminds the traditional Agent if intention (volition) is a feature of agentivity[64].
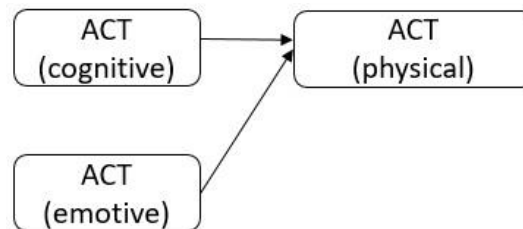
We have the following causal relation (see Fig. 18):

Figure 18. Causal relation between ACT (cognitive), ACT (emotive) and ACT (physical).

Example-1: the sentence *After deep and thorough thinking, John pressed the button* represents the causal relation between *ACT (cognitive)* and *ACT (physical).*

Example-2: the sentence *Being scared John hit Peter* represents the causality between *ACT (emotive)* and *ACT (physical);*

---

[63] Schank used the primitive *MBUILD* instead of *ACT (cognitive)* to create or combine thoughts.

[64] Although there are some exclusions like verbs *kill* and *murder*, English does not have an explicit method of marking lexical categories for volition or non-volition. When English speakers want to be clear about whether an action was done intentionally or not, adverbs such as "intentionally", or "accidentally" are included in the sentence (Wikipedia "Volition").

There are also causal relations between *ACT (Cognitive)* and *ACT (emotive)* (Fig. 19):
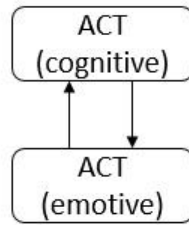

Figure 19. Causal relations between ACT (cognitive) and ACT (emotive).

Example-1: the sentence *My fear made me think over his behavior* represents causal relation from *ACT (emotive)* to *ACT (cognitive).*
Example-2: the sentence *What I knew made me laugh* represents causal relation from *ACT (cognitive)* to *ACT (emotive).*

The primitive *PERCEIVE*
Now we look at both *ACT (cognitive)* and *ACT (emotive)* to consider what primitive can cause them. The trigger for the activation of these primitives is information that an animate object gets from the outside world. Information is received via sense organs, so the primitive causing either *ACT (cognitive)* or *ACT (emotive)* is the primitive *PERCEIVE*[65]. It represents the class of verbs related to receiving sensory information: *hear, see, smell,* etc.
Checking the criteria:

a) On the conceptual level, there are two semantic roles for the primitive *PERCEIVE* that correspond to two positions: *who perceives* or *Perceiver* and *perceive what* or *Stimulus*[66].

b) On the syntactic level, the primitive can be:
- ➤ asymmetrical and symmetrical. If *X sees Y* it is possible to have both asymmetrical and symmetrical cases: *Y doesn't see X* and *Y sees X.*
- ➤ reflexive: *X sees his hand.*
- ➤ unary relation (optional): unknown.

We have the following causal relation (see Fig. 20):

---

[65] Instead of the primitive *PERCEIVE,* Schank used the primitive MTRANS: "to transfer information mentally".

[66] In (Engelkamp and Zimmer, 2006) the following sequence has been extracted: stimulus—retinal image—early processing (features)—structural description—matching stored structural description—retrieval of stored functional knowledge. Roughly, one can consider the process "stimulus—retinal image" as *PERCEIVE* while the rest as *ACT (cognitive).*
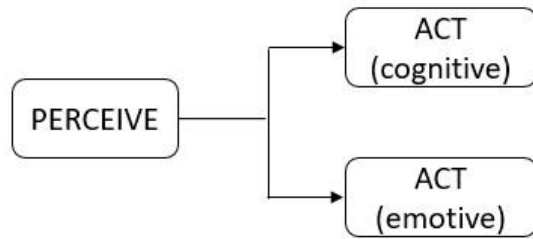
Figure 20. Causal relation between *Perceive, ACT (cognitive)* and *ACT (emotive).*

Example-1: the sentence *Dark spot on my X-ray frightened me* means that I saw (perceived) the dark spot and it causes *ACT (emotive)* that led to bad feelings.
Example-2: the sentence *Dark spot on my X-ray forced me think about it* means that I saw (perceived) the dark spot and it causes *ACT (cognitive)*.
Example-3: the sentence *Dark spot on my X-ray scared me and forced to think* combines *PERCEIVE* with both *ACT (cognitive)* and *ACT (emotive).*

The primitives *Perceive, ACT (cognitive)* and *ACT (emotive)* allow us to see the traditional role of Experiencer as being too wide and mixed. Experiencer combines perception and reaction on it. In other words, *Experience* as a primitive does not separate multi-stage processes from perception to cognitive/emotional processes. For example, the sentence *I see sunset* means *I get the image of sunset in my memory (PERCEIVE)* while the sentence *I like sunset* means *I like the image of sunset in my memory (ACT (emotive))*. Finally, the sentence *I see the beautiful sunset* means the causality between *PERCEIVE* and *ACT (emotive).*

Primitive *ACT (informational)*
As mentioned, the primitives *ACT (cognitive)* and *ACT (emotive)* can make animate objects implement *ACT (physical)*. They can also make animate objects implement *ACT (informational).* In English, informational verbs are usually triadic and include the traditional Agent, Topic and Recipient -- *X told about Y to Z* -- which means that the process of information transfer assumes the presence of transmitter (Agent), information (Topic) and receiver (Recipient). But in reality *transmit* doesn't assume *receive*. *ACT (informational)* purely represents a process of information transmission that should be completed by *PERCEIVE* as a process of receiving information (*I listened carefully to what he was saying to me*). The *Actor (informational)* may have a particular recipient in mind but this is not the point for *ACT (informational)*. Separation of *transmit* from *receive* allows to explain confusion in role assignment: in the sentence *John listens to music, John* is Experiencer but in *John listens to Peter, John* is Patient.

   ACT (informational) represents the class of verbs related to information transfer: *inform about, report about, instruct about, talk about, speak about, write about,* etc.
Checking the criteria:

a) On the conceptual level, there are two semantic roles for the primitive *ACT (informational)* that correspond to two positions: *who act informationally* or *Actor (informational)* and *act what* or *Actee (informational)*. There is an important difference between *Actee* for *ACT (physical)* and *ACT (informational).* While *Actee* for *ACT (physical)* is an object affected by a physical act, *Actee* for *ACT (informational)* is a text either spoken (speech) or written (letter).

b) On the syntactic level, the primitive can be:
   ➢ asymmetrical and symmetrical. If *X talks about Y*, it is possible to have both *Y does not talk about X* and *Y talks about X*.
   ➢ reflexive: *X talks about X.*
   ➢ unary relation (optional). *ACT (informational)* is unary if it is made spontaneous: *X groan* (*shout, sigh,* etc).

We have the following causal relation (see Fig. 21):
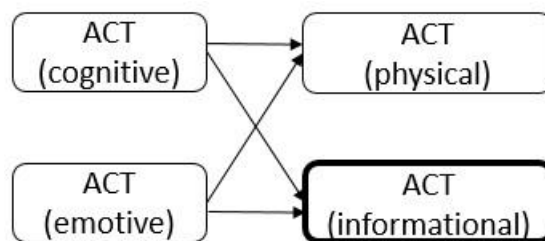
Figure 21. Causal relation between ACT (cognitive), ACT (emotive) and ACT (informational).

Example-1: the sentence *Becoming calmer he wrote the letter* represents the causality between *ACT (emotive)* and *ACT (informational);*
Example-2: the sentence *After thinking about the project, he made a presentation about tsunami* represents the causality between *ACT (cognitive)* and *ACT (informational);*

Conclusion
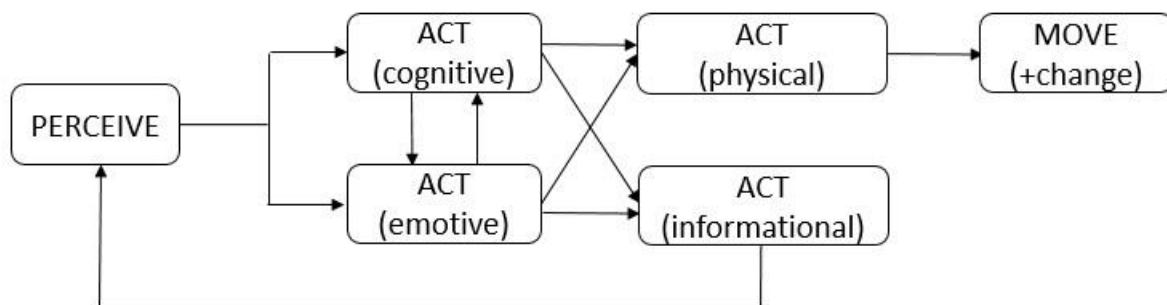The entire set of primitives linked by causal relation is shown in Fig. 22:

Figure 22. Primitives with causal relations.

Example: the sentence *Colonel marched the soldiers to their tents* on the conceptual level will be unfolded with all primitives presented in Fig. 22:

> *(colonel ACT (informational) soldiers) CAUSE*
> *(soldiers PERCEIVE command) CAUSE*
> *(command ACT (cognitive and emotive) soldiers) CAUSE*
> *(soldiers ACT (physical) soldiers) CAUSE*
> *(soldiers MOVE tents)*

It is common to have gaps in description on the syntactic level. In the sentence *He read the note and shot himself*, some elements are hidden and need to be restored to get the full chain of causal relations between primitives.

4.6. Semantic role on the conceptual level

In light of the investigation in 4.1.-4.5., we will:

➢ revise the semantic role definition;

➢ answer the two issues formulated in the Introduction (ch.1):

   Role inventory: how many semantic roles are needed and what they are.

   Role combination: how to distinguish a valid set of roles from an invalid one.

4.6.1. Revision of semantic role definition

A semantic role is defined as a conceptual relation between a position and a primitive:

$$SR_{p_i}^{Prim_j} = R^C(p_i, Prim_j)$$

In this case, the relation between roles that determines their combination (2nd issue) will be defined as a relation between relations R^C, which, in turn, contain a primitive as a relation. It appears complicated.

Relation R^C can be clarified. Following 4.3., there are two and only two positions for any primitive. The difference between them is their location with respect to a primitive (or a direction towards a primitive). Relation R^C can be defined in a simple way as a concatenation of a position and a primitive.

Definition of concatenation. For any primitive *Prim*, a concatenation of a position and a primitive is defined as a functional feature of an entity that either initiates a primitive (*who/what Prim)* or is affected by a primitive (*Prim on whom/what):*

(8) $R^C(p_i, Prim_j) = (p_1 {}^\wedge Prim_j) \veebar (Prim_j {}^\wedge p_2)$

where ^ is a sign of concatenation; $\veebar$ is exclusive disjunction.

Since concatenation is not commutative, i.e., ab ≠ ba, there is no need to distinguish positions p_1 and p_2 by lowercase digits. Concatenation clearly indicates which position is involved in it. So, (8) can be simplified as follows:

$$R^C\left(p, Prim_j\right) = (p\wedge Prim_j) \veebar (Prim_j\wedge p)$$

Definition of a semantic role. A semantic role is a concatenation of a position and a primitive:

$$SR_p^{Prim_j} = (p\wedge Prim_j) \veebar (Prim_j\wedge p)$$

Since each primitive has two positions, semantic roles are *pairwise entities*[67].
For example, the primitive *ACT* (all types) produces roles *who/what ACT* (=*Actor*) and *ACT on whom/what* (=*Actee*)[68]; the primitive *PERCEIVE* produces roles *who PERCEIVE* (=*Perceiver*) and *PERCEIVE whom/what* (=*Stimulus);* the primitive *MOVE* produces roles *what MOVE* and *MOVE with respect to what,* etc*.*

4.6.2. Role inventory: how many semantic roles are needed and what they are.
The number of roles equals the number of positions for all extracted primitives (or multiplication of the number of primitives by two)[69]. Having six extracted primitives, the number of roles equals 12 (see Table 3).

4.6.3. Role combination: how to distinguish a valid set of roles from an invalid one.
Semantic roles are related by their nature as being pairwise entities (see 4.6.1.). Role twoness is the base for determination of a valid set of roles since twoness prevents the other roles from being involved in a primitive. A minimal valid set of roles consists of pairwise roles linked to one primitive. In the case of the causal relation between *n* primitives (see 4.5.), a valid set of roles consists of *n* pairwise roles accordingly.

---

[67] The term "paired semantic roles" has already been used in (Perini, 2015:189) in another context. Perini considers semantic roles only on syntactic level and, according to him, not all semantic roles participate in a pairing relationship.

[68] Actually, it is a matter of terminology on how to name them. The easiest way is to follow Propbank's approach where Agent for the verb *hit* is named as *hitter.*

[69] From one point of view, this is the answer to the question James Pustejovsky raised in (1990): What is the computational and representational trade-off between fixed semantic roles and decomposition into primitives?

Primitives	Semantic roles
PERCEIVE	Perceiver
	Stimulus
ACT (cognitive)	Actor (cognitive)
	Actee (cognitive)
ACT (emotive)	Actor (emotive)
	Actee (emotive)
ACT (physical)	Actor (physical)
	Actee (physical)
ACT (informational)	Actor (informational)
	Actee (informational)
MOVE	Moving object
	Object relative to which movement occurs

Table 3. List of primitives and semantic roles.

5. Conclusion

Investigating two fundamental issues in the semantic role theory, we arrived at the following key statements:

1. Semantic roles are elements of the conceptual level of a language, not the syntactic level, which means that a semantic role is not defined in terms of arguments and a verb.

2. On the conceptual level, a semantic role is defined via a conceptual structure that consists of positions and a primitive.

3. Positions on the conceptual level in comparison with positions on the syntactic level are mandatory elements. There two and only two positions for each primitive.

4. To be role-represented, a primitive must (a) have two positions on the conceptual level; and (b) be a binary relation which is reflexive, asymmetrical and symmetrical at once on the syntactic level (optionally, it can be a unary relation as well). Not all primitives and not all verbs accordingly are role-represented, which means that semantic roles are not a universal tool for semantic representation.

5. Role-represented primitives can be extracted. The procedure is based on the criteria (a) & (b) from 4. and on the assumption that each primitive has at least one causal relation with another one. In total, six primitives were extracted.

6. A semantic role is defined as a feature defined by the concatenation of a position and a primitive.

7. The number of roles equals the number of positions for all extracted primitives (or the number of primitives multiplied by two). *This conclusion answers the issue about role inventory:* the number of roles is 12 and they are defined based on the extracted primitives.

8. Semantic roles are pairwise entities. *This conclusion answers the issue about role combination:* a minimal valid set of roles consists of pairwise roles linked to one primitive. In the case of a causal relation between *n* primitives, a valid set of roles consists of *n* pairwise roles accordingly.

Acknowledgements

Appendix: comparison of semantic roles on the syntactic and conceptual level

Semantic roles on the syntactic level from the Project LIRICS (Linguistic InfRastructure for Interoperable ResourCes and Systems)[70] will be compared with semantic roles on the conceptual level. Project LIRICS was chosen since its set of roles was used as the base for developing a standard set of semantic roles for the International Organization for Standardization[71].

Table 4 demonstrates the comparison. The content of columns 1-3 (role, definition, example) are taken from LIRICS. In column 4 (status) there are five possibilities regarding the status of LIRICS role on the conceptual level:

1. E*xact match:* LIRICS role coincides with a role on the conceptual level;

2. *Partial match:* LIRICS role is broader than a role on the conceptual level;

3. *Particular case*: LIRICS role is a particular case of a role on the conceptual level;

4. *Derived role*: LIRICS role is a compound of two roles on the conceptual level;

5. *Not a role.* LIRICS role is not a role on the conceptual level. There are several reasons for that (see column 5 for details):

> --statement of completeness, i.e. obligation to assign a role to every argument for every verb;
> --IS-A relation (for details see 2.3.1.1. Range of arguments);
> --role is perspective-dependent (for details see 2.3.1.1. Range of verbs);
> --role represents an entire event, not a participant, etc.

---

[70] https://lirics.loria.fr/doc_pub/D4-3.pdf

[71] https://www.iso.org/obp/ui/#iso:std:iso:24617:-4:ed-1:v1:en

Role in LIRICS	Definition in LIRICS	Example	Status of LIRICS role on conceptual level	Comments
Agent	Participant in an event who is animate or perceived as animate, who initiates and carries out the event intentionally or consciously, and who exists independently of the event.	*"John [agent] built the house"*	Derived role	Intention assumes causal relation between two primitives, for example, *ACT (cognitive) CAUSE ACT (physical)*
Partner	Participant in an event who is animate or perceived as animate, who is not the principal agent of the event, but who is intentionally or consciously involved in carrying out the event, and who exists independently of the event.	*"John [agent] built the house with Stephen [partner]"*	Not a role	Statement of completeness forces us to set Partner. On the conceptual level, Partner merges with Agent : "John with Stephen [agent] built the house"
Cause	Participant in an event (that may be animate or inanimate) that initiates the event, but that does not act with any intentionality or consciousness; it exists independently of the event.	*"The bomb [cause] started several secondary fires"*	Exact match	Cause = Actor
Instrument	Participant in an event that is manipulated by an agent, and with which an intentional act is performed.	*"He opened the door with the key [instrument]"*	Derived role	see 4.3.4.1 for details
Patient	Participant in an event that undergoes a change of state, location of condition, that is causally involved or directly affected by other participants, and exists independently of the event.	*"John hit Mary [patient]"*	Exact match	Patient = Actee
Pivot	Participant in a state that has a major or central role or effect in that state. A pivot is characterized as being in a certain position or condition throughout the state, and is more central to the state than a participant in theme role.	*"John [pivot] owns that dog [theme]"*	Not a role	Pivot is a perspective-dependent role (see 2.3.1.1. Range of verbs): *John owns that dog* vs. *That dog belongs to John*
Theme	Participant in a state or event that is essential to the event taking place or the state being in effect. Theme is distinguished from the semantic role of patient principally in that it is not affected or changed by the event.	*"John owns that dog [theme]"* *"He talked about politics [theme]"*	Partial match	Actee (informational)
Beneficiary	Participant in a state or an event that is advantaged or disadvantaged by the event or state.	*"John sold the car for a friend [beneficiary]"* *"He gave his life for his country [beneficiary]"*	Partial match	Object relative to which movement occurs

Source	Participant in an event that is the (non-locative, non-temporal) start point of an action. The source exists independently of the event.	*"The researcher got his idea from a book [source]"*	Particular case	Source is a particular case of "object relative to which movement occurs"
Goal	Participant in an event that is the (non-locative, non-temporal) end point of an action. The goal exists independently of the event.	*"Edison customers [goal] have received electric services since April"*	Particular case	Goal is a particular case of "object relative to which movement occurs"
Result	Participant in an event that is an inanimate entity (or entities) that describes a terminal point which will be reached in the normal course of events or in all possible courses of events.	*"She built a house [result]"*	Exact match	Object relative to which movement occurs
Reason	Participant that represents the set of facts or circumstances explaining why a state exists or an event occurs. In other words, the reason is the source of the state or the cause of the event.	*"People love giant pandas because they look like teddy bears [reason]"*	Not a role	Reason represents the entire event that causes another one.
Purpose	Participant that represents the set of facts or circumstances that describe what an agent wishes or intends to accomplish by performing some intentional action.	*"The robber tied Harry to the chair to stop him from getting away [purpose]*	Not a role	Purpose represents the entire event that causes another one.
Time	Participant that indicates an instant or an interval of time during which a state exists or an event took place.	*"Dinner is at 6 o'clock [time]"*	Not a role	IS-A relation
Manner	Participant that represents the way or style of performing an action or the degree/strength of the cognitive state or perception.	*"Lester was coldly [manner] polite."* *"The tiny stick was fastened tightly [manner] to his wrist."*	Not a role	IS-A relation
Medium	Participant that represents the physical or abstract setting, entity or channel used by an agent or agents in an event or process.	*"The students heard the news on the radio [medium]"*	Derived role	see 4.3.4.1 for details
Means	Participant that represents an abstract entity (or entities) or a procedure of the action in terms of component steps of an event. The means is the method by which an intentional act is performed by an agent. The role means differs from instrument in that means describes abstract things (abstract means and methods) while the instrument describes concrete things.	*"The mayor delayed the ribbon ceremony by pretending to be ill [means]"* *"He had to button his sleeve by holding the cuff in his mouth [means]"*	Not a role	Means represents the entire event that is related to or causes another one.
Setting	Participant that represents the abstract setting for the occurrence of an event, or a state, or a fact.	*"Libya employed chemical weapons in the conflict [setting]"*	Not a role	Setting is a property of event itself (similar to Location).

Location	Participant that represents the place where an event occurs, or a state that is true, or a thing that exists.	*"She was cooking in the kitchen [location]"*	Not a role	Location is either a property of an event itself or perspective-dependent role (see 2.3.1.1. Range of verbs)
Initial Location	Participant that indicates the place where an event begins or a state becomes true.	*"Half way out of the harbor [initial Location] the sea becomes really deep"*	Not a role	Location is either a property of an event itself or perspective-dependent role (see 2.3.1.1. Range of verbs)
Final Location	Participant that indicates a place where an event ends or a state becomes false.	*"The race finishes in Tilburg [final Location]"*	Not a role	Location is either a property of an event itself or perspective-dependent role (see 2.3.1.1. Range of verbs)
Path	Participant that indicates an intermediate place or state or trajectory between two locations, or in a designated space.	*"The two men climbed the slippery slope [path]"*	Particular case	Path is a particular case of "object relative to which movement occurs"
Distance	Length or extent of space.	*"Terry jogged for two miles [distance]"*	Not a role	IS-A relation
Amount	Quantity denoting participant.	*"I have several [amount] brothers"*	Not a role	IS-A relation
Attribute	Property of an entity or entities.	*"Roses are red [attribute]"*	Not a role	IS-A relation
Frequency	Number of occurrences within a given time span.	*"He washed the car religiously twice a week [frequency]".*	Not a role	IS-A relation

Table 4. Comparison of LIRICS roles with roles on the conceptual level.

References

Abend, Omri, & Ari Rappoport (2017), The State of the Art in Semantic Representation. *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*. Vancouver, Canada.

Ackerman, F., & Moore, J. (1998), *Syntagmatic and Paradigmatic Dimensions of Causee Encodings*. Available on the LFG Website.

Ackerman, F. & Moore, J. (2001), *Proto-properties and Grammatical Encoding: A Correspondence Theory of Argument Selection*. Stanford: CSLI.

Alexiadou, Artemis & Schäfer, Florian (2006), Instrument Subjects Are Agents or Causers. *Proceedings of the 25th West Coast Conference on Formal Linguistics*. Cascadilla Proceedings Project, Somerville, MA. http://www.lingref.com/cpp/wccfl/25/paper1431.pdf

Allen, James & Teng, Choh Man (2018), Putting Semantics into Semantic Roles. In *Proceedings of the Seventh Joint Conference on Lexical and Computational Semantics*. New Orleans, Louisiana. 235—244.

Anderson, John (1971), *The Grammar of Case: Towards a Localistic Theory*. Cambridge: Cambridge University Press.

Baker, Mark (1997), *Thematic Roles and Syntactic Structure*. Ms. McGill University.

Banarescu, Laura, Claire Bonial, Shu Cai, Madalina Georgescu, Kira Griffitt, Ulf Hermjakob, Kevin Knight, Philipp Koehn, Martha Palmer, & Nathan Schneider (2013), Abstract meaning representation for sembanking. In *Proceedings of the 7th Linguistic Annotation Workshop and Interoperability with Discourse*. 178-186.

Begum, Gousiya, & Sultana, N. Musrat (2013), Concept Based Mining Model for Text Clustering. In *International Conference on Advances in Computer Science and Engineering (ICACSE)*, Volume2, No.1.: 154-158.

Bonial, Claire, William Corvey, Martha Palmer, Volha V. Petukhova, & Harry Bunt (2011), A Hierarchical Unification of LIRICS and VerbNet Semantic Roles. In *Proceedings IEEE-ICSC 2011 Workshop on Semantic Annotation for Computational Linguistic Resources*. 1-7.

Bresnan, Joan (1982), The Passive in Lexical Theory. In Joan Bresnan (ed.), *The Mental Representation of Grammatical Relations*, MIT Press, Cambridge. 3–86.

Bunt, Harry, & Martha Palmer (2013), Conceptual and representational choices in defining an ISO standard for semantic role annotation, In *Proceedings of the ACL-ISO Workshop on Interoperable Semantic Annotation (ISA-9)* held in conjunction with *International Workshop on Computational Semantics*, Potsdam, Germany, March, 2013.

Carlson, Greg (1998), Thematic Roles and the Individuation of Events. In *Events and Grammar* (ed. Susan Rothstein), Kluwer, Dordrecht. 35–51.

Chia-Hung Lin, Chia-Wei Yen, Jen-Shin Hong, & Samuel Cruz-Lara (2007), Event-based textual document retrieval by using semantic role labeling and coreference resolution. In *IADIS International Conference WWW/Internet*. Vila Real, Portugal.

Chunhua Weng, Xiaoying Wu, Zhihui Luo, Mary Regina Boland, Dimitri Theodoratos, & Stephen B Johnson (2011), EliXR: an approach to eligibility criteria extraction and representation. *Journal of the American Medical Informatics Association*, Volume 18, Issue Supplement 1: i116–i124.

Croft, William (1998), Event structure in argument linking. In Miriam Butt and Wilhelm Geuder (eds.) *The projection of arguments: lexical and compositional factors.* Stanford: Center for the Study of Language and Information. 1-43.

Croft, William (2012), *Verbs: Aspect and causal structure*. Oxford University Press.

Culicover, P.W., & Wilkins, W. (1986), Control, PRO, and the Projection Principle. *Language* 62(1): 120-153.

Davis, Anthony & Koenig, Jean-Pierre (2000), Linking as Constraints on Word Classes in a Hierarchical Lexicon. *Language.* 76: 56-91.

DeLancey, Scott (1991), Event Construal and Case Role Assignment. *Proceedings of the Seventeenth Annual Meeting of the Berkeley Linguistics Society: General Session and Parasession on The Grammar of Event Structure.* 338-353

DLPP=*Deep Learning with PaddlePaddle*. At:
https://www.paddlepaddle.org.cn/documentation/docs/en/1.3/beginners_guide/basics/label_semantic_roles/index_en.html

Dowty, David (1989), On the semantic content of the notion of "thematic role". In G. Chiercha, B. H. Partee, & R. Turner (Eds.), *Properties, types and meaning*. Dordrecht: Kluwer. 69-129.

Dowty, David (1991), Thematic Proto-Roles and Argument Selection. *Language*, 67 (3): 547-619.

Eckardt, Regine (2013), *Adverbs, Events, and Other Things: Issues in the Semantics of Manner Adverbs.* Walter de Gruyter.

Engelkamp, Johannes & Zimmer, Hubert D. (2006), *Lehrbuch der Kognitiven Psychologie*. Gottingen, Hogrefe.

Grimshaw, J. (1993), Semantic structure and semantic content in lexical representation. In J. Grimshaw (ed.) (2005) *Words and structure*, Stanford: CSLI. 75-89.

Gruber, Jeffrey S. (1965), *Studies in Lexical Relations*. Doctoral Dissertation, MIT.

Faulhaber, S. (2011), *Verb Valency Patterns. A Challenge for Semantics-Based Accounts*. Berlin, Boston: De Gruyter Mouton.

Fillmore, Charles J. (1968), The Case for Case. In Bach & Harms (Ed.): *Universals in Linguistic Theory*. New York: Holt, Rinehart, & Winston. 1-88.

Fillmore, Charles J. (1971), Some Problems for Case Grammar. In *Working Papers in Linguistics*, no. 10 (1971), Ohio State University. Department of Linguistics. 245-265.

Fillmore, Charles J. (1977), The case for case reopened. In *Syntax and Semantics, Vol. 8: Grammatical   Relations.* New York: Academic Press Inc.

Fillmore, Charles J., Baker Collin F. & Sato Hiroaki (2002), The framenet database and software tools. In *Proceedings of the Third International Conference on Language Resources and Evaluation (LRE).*

Fillmore, Charles J. & Kay Paul (1993), *Construction Grammar.* Ms., U.C. Berkeley.

Finegan, Edward (2007), *Language: Its Structure and Use*. Cengage Learning.

Fried, Mirjam & Ostman, Jan-Ola (2004), *Construction Grammar in a Cross-language Perspective.* John Benjamins Publishing.

Ellsworth, Michael, Erk, Katrin, Kingsbury, Paul & Pado, Sebastian (2004), PropBank, SALSA, and FrameNet: How Design Determines Product. In *Proceedings of the LREC-2004. Workshop on Building Lexical Resources from Semantically Annotated Corpora*. Lisbon.

Hanxiao Shi, Guodong Zhou & Peide Qian (2010), An Attribute-based Sentiment Analysis System. *Information Technology Journal,* 9: 1607-1614.

Haspelmath, Martin (2006), Terminology of case. In A. Malchukov & A. Spencer (eds.), *Handbook of Case*, Oxford University Press.

Huminski, A. & Zhang H. (2018), *Event Representation through Semantic Roles: Evaluation of Coverage*. arXiv:1810.03879.

Huminski, A., Zhang, H. & Krishnamurthy G. (2018), *Towards Verifying Semantic Roles Co-occurrence.* arXiv:1810.03875.

Huminski, A., Liausvia, F. & Goel, A. (2019), Semantic Roles in VerbNet and FrameNet: Statistical Analysis and Evaluation. In *20th International Conference on Computational Linguistics and Intelligent Text Processing (CICLing-2019).* La Rochelle (France). Apr. 07-13.

Jackendoff, Ray (1972), *Semantic Interpretation in Generative Grammar*. Cambridge, MA: MIT Press.

Jackendoff, Ray (1983), *Semantics and cognition*. Cambridge, MA: MIT Press.

Jackendoff, Ray (1987), The Status of Thematic Relations in Linguistic Theory. *Linguistic Inquiry* 18 (3): 369-411

Jackendoff, Ray (1990), *Semantic Structures*, MIT Press, Cambridge, MA.

Jackendoff, Ray (1993), The combinatorial structure of thought: The family of causative concepts. In E. Reuland, and W. Abraham (eds.), *Knowledge and Language: Volume II. Lexical and Conceptual Structure*. 31-49.

Jackendoff, Ray (2007), A whole lot of challenges for linguistics. *Journal of English Linguistics*, 35:253-62.

Jurafsky, Dan & Martin, James (2008), *Speech and Language Processing.* Upper Saddle River, NJ, United States.

Kamp, Hans & Rossdeutscher, Antje (1994), Remarks on lexical structure and DRS construction. *Theoretical Linguistics,* 20:97-164.

Kasper, Simon (2008), *A comparison of "thematic role" theories.* Thesis for Magister Artium. Heinrich-Heine-Universität Düsseldorf.

Kasper, Simon (2010), *What is an agent supposed to be?* Philipps-Universität Marburg. Marburg.

Kipper Schuler, Karin (2005), *VerbNet: A broad-coverage, comprehensive verb lexicon.* Ph.D. thesis. Computer and Information Science Dept. University of Pennsylvania. Philadelphia. PA.

Kittila, Seppo, Vasti, Katja & Ylikoski, Jussi (2011), Introduction to case, animacy and semantic roles. In: *Case, Animacy and Semantic Roles*, Amsterdam: John Benjamins. 1-26.

Kittila, Seppo & Zuniga, Fernando (2016), Recent development and open questions in the field of semantic roles. In *Advances in Research on Semantic Roles.* John Benjamins Publishing Company, Amsterdam/Philadelphia.

Koenig, Jean-Pierre & Davis, Anthony R. (2001), Sublexical Modality and the Structure of Lexical Semantic Representations. *Linguistics and Philosophy.* 24:71–124.

Lang, Joel & Mirella Lapata (2014), Similarity-Driven Semantic Role Induction via Graph Partitioning. *Computational Linguistics*, 40 (3).

Le-Hong Phuong, Pham Thai Hoang, Pham Xuan Khoai, Nguyen Thi Minh Huyen, Nguyen Thi Luong, & Nguyen Minh Hiep (2017), Vietnamese Semantic Role Labelling. *VNU Journal of Science: Comp. Science & Com. Eng.*, 33(2):39-58.

Levin, Beth (1995), Approaches to Lexical Semantic Representation, in D. Walker, A. Zampolli, & N. Calzolari, eds., *Automating the Lexicon I: Research and Practice in a Multilingual Environment,* Oxford University Press, Oxford. 53-91.

Levin, Beth (1999), Objecthood: An Event Structure Perspective, *Proceedings of CLS 35, volume 1: The Main Session,* Chicago Linguistic Society, University of Chicago, Chicago, IL. 223-247.

Levin, Beth, & Malka Rappaport Hovav (2005), *Argument Realization*, Research Surveys in Linguistics Series, Cambridge University Press, Cambridge, UK

Levin, Beth (2014), Semantic Roles. In M. Aronoff, ed., *Oxford Bibliographies in Linguistics*, Oxford University Press, New York.

Levin, Beth (2019), On Dowty's `Proto-roles and Argument Selection. At: http://web.stanford.edu/~bclevin/dowty19fin.pdf To be appeared in L. McNally & Z. Szabó, eds., *A Reader's Guide to Classic Papers in Formal Semantics*, Studies in Linguistics and Philosophy 100, Springer, Dordrecht.

Lin, Jimmy (2004), *Event structure and the encoding of arguments: the syntax of the Mandarin and English verb phrase. Ph.D. dissertation. Massachusetts Institute of Technology, Cambridge, MA, USA.*

LIRICS (2007), Deliverable D4.3. At https://lirics.loria.fr/doc_pub/D4-3.pdf.

Mefteh, Mariem, Bouassida, Nadia & Hanêne, Ben-Abdallah (2016), Mining Feature Models from Functional Requirements. *The Computer Journal*, 59(12):1784–1804

Majid Hadhir, Mohammed & Ali, Hasan Shaban (2019), Semantic Roles in Arabic with Reference to English. *Journal of Al-Frahedis Arts.* 37(1):510-523

Morante, Roser (2008), Semantic Role Labeling Tools Trained on the Cast3LB-CoNNL-SemRol Corpus in *Proceedings of the Sixth International Conference on Language Resources and Evaluation* (LREC'08). Marrakech, Morocco.

Moreda P., Palomar M. & Suarez A. (2004), SemRol: Recognition of Semantic Roles. In *Advances in Natural Language Processing: 4th International Conference*, EsTAL 2004, Alicante, Spain.

Næss, Ashild (2007), *Prototypical Transitivity*. J. Benjamins Pub.

Newmeyer, Frederick J. (2010), On comparative concepts and descriptive categories: A reply to Haspelmath. *Language* 86(3):688-695

OBO = Oxford Bibliographies. (2014), Semantic Roles. At:
<https://www.oxfordbibliographies.com/view/document/obo-9780199772810/obo-9780199772810-0141.xml>.

Ochoa, José Luis, Hernández-Alcaraz Maria Luisa, Valencia-García Rafael & Martínez-Béjar Rodrigo (2011), A semantic role-based methodology for knowledge acquisition from Spanish documents. In *International Journal of the Physical Sciences.* 6(7):1755-1765.

Palmer, Martha, Daniel Gildea, & Paul Kingsbury (2005), The proposition bank: An annotated corpus of semantic roles. *Computational linguistics* 31(1):71–106

Payne, Thomas E. (1997), *Describing morphosyntax: A guide for field linguists.* Cambridge; New York: Cambridge University Press.

Carreras, Perez Xavier (2005), *Learning and Inference in Phrase Recognition: A Filtering-Ranking Architecture using Perceptron*. PhD thesis, Universitat Politecnica de Catalunya. Barcelona.

Perini, Mario A. (2015), *Describing Verb Valency: Practical and Theoretical Issues.* Springer International Publishing, Switzerland.

Petukhova, Volha, & Harry Bunt (2008), LIRICS Semantic Role Annotation: Design and Evaluation of a Set of Data Categories. In *Proceedings of the Sixth International Conference on Language Resources and Evaluation (LREC)*, Marrakech, Morocco.

Pollard, Carl, & Ivan Sag (1994), *Head-Driven Phrase Structure Grammar*. Stanford: Center for the Study of Language and Information.

Primus, Beatrice (1999), *Cases and Thematic Roles: Ergative, Accusative and Active*. Niemeyer, Tubingen.

Primus, Beatrice (2006), Mismatches in Semantic-Role Hierarchies and the Dimension of Role Semantics. In I. Bornkessel et al. (eds.) *Semantic Role Universals and Argument Linking. Theoretical, Typological and Psycho-/Neurolinguistic Perspectives.* Berlin: Mouton de Gruyter. 53-87.

Priya, K. V., & Mathew Kurian (2011), Mining Event-based Commonsense Knowledge from Web using NLP Techniques. *IJCA Special Issue on Computational Science - New Dimensions & Perspectives* (1):9-12.

Pustejovsky, James (1990), Guest Editor's Note: On the Nature of Lexical Knowledge. *Machine Translation,* 5(1):1-4.

Rappaport, Malka, & Beth Levin (1988), What to Do with Theta-Roles. In W. Wilkins, ed., *Syntax and Semantics 21: Thematic Relations,* Academic Press, New York, NY. 7-36.

Riemer, N. (2010), *Introducing Semantics.* New York: Cambridge University Press.

Saeed, John I. (1997), *Semantics*. Oxford: Blackwell Publishers.

Sánchez, Liliana & Oliveira, Claudia (2008), Deverbal noun complementation rules applied to semantic role labeling. *Revista da ABRALIN*. 7(2):9-42.

Schank, Roger (1973a), *The fourteen primitive actions and their inferences*. SUAIM-83.

Schank, Roger (1973b), Identification of Conceptualizations Underlying Natural Language. In Roger C. Schank and Kenneth Mark Colby (eds.), *Computer Models of Thought and Language.* W. H. Freeman & Co., San Francisco.

Schank, Roger (ed.). (1975), *Conceptual Information Processing*. Amsterdam: North-Holland.

Schank, Roger, & Riesbeck, C. K. (1981), Inside Computer Understanding: Five Programs Plus Miniatures. In *Artificial Intelligence Series.*

Schlesinger, I. (1989), Instruments as Agents: On the Nature of Semantic Relations. *Journal of Linguistics, 25*(1):189-210.

Schiffrin, A. & Bunt H.C. (2007), LIRICS Deliverable D4.3 Documented compilation of semantic data categories. In: https://lirics.loria.fr/doc_pub/D4-3.pdf

Somers, Harold L. (1987), *Valency and Case in Computational Linguistics.* Edinburgh University Press, Edinburgh, Scotland.

Stalmaszczyk, P. (2007), *Theta roles and the theory of theta-binding.* University of Lodz.

Sugisaki, Kyoko (2016), *Automatic Annotation and Assessment of Syntactic Structures in Law Texts Combining Rule-Based and Statistical Methods*. PhD thesis. University of Zurich.

Talmy, Leonard (1985), Force Dynamics in Language and Thought. In *Papers from the 21st Regional Meeting*. *Chicago Linguistic Society.* University of Chicago. Chicago. Illinois.

Valdimarsson, Sigurbjörn Már (2012), *Semantic Roles - The Aspectual Interface Hypothesis and Argument Realization*. At https://skemman.is/bitstream/1946/12867/1/Semantic%20Roles%20-%20The%20Aspectual%20Interface%20Hypothesis%20and%20Argument%20Realization.pdf

Van Valin, Robert D. (1999), Generalized Semantic Roles and the Syntax-Semantics Interface. In F. Corblin, C. Dobrovie-Sorin, & J.-M. Marandin (Eds.), *Empirical issues in formal syntax and semantics*.

Van Valin, Robert, & Randy LaPolla (1997), *Syntax. Structure, Meaning, and Function*. Cambridge: Cambridge University Press.

Yule, George (2006), *The Study of Language*. Cambridge, UK: Cambridge University Press.