# Valency and Semantic Features of Verbs

V. Benešová

Charles University, Faculty of Mathematics and Physics, Prague, Czech Republic.

**Abstract.**  In the contribution, different approaches to verb classification are characterized. We briefly describe valency theory of Functional Generative Description that provides the theoretical background for valency lexicon VALLEX. The basic criteria for classifying verbs from the valency lexicon VALLEX are formulated. Some illustrative instances of delimitation of classes such as communication, mental action, perception, exchange and providing are introduced in this paper.

## 1. Introduction

The classification of verbs into semantic classes appears to be an "up-to-date" question in linguistic research at present. Sorting verbs could seem to be self effective but it is important to emphasize that verb classes can considerably contribute to checking the data consistency because such a classification allows to observe the whole groups of semantically similar verbs together. Furthermore, the measure of successfulness can reveal strengths of the used descriptive instruments and as a result, it makes possible to draw attention to weak or peripheral places in the description. In the present contribution, we first compare some of the existing approaches to the valency of verbs, namely Levin's verb classes, PropBank, LCS, FrameNet, and classifications based on these approaches and we present the valency approach of the Functional Generative Description and the valency lexicon VALLEX. Our main objective is to attempt at a formulation of criteria for the classification of verbs based on the mentioned framework and to outline a preliminary possible classification arising from data in VALLEX.

## 2. Some Approaches to the Classification of Verbs

We briefly characterize valency structures designed for English in **Levin´s classification, PropBank, Lexical Conceptual Structure** (**LCS**) and **FrameNet.**

**2.1** Let us start with **Levin´s verb classes** [Levin, 1993]**.** Levin observes syntactic properties of a high number of English verbs and points out that some aspects of the syntactic behavior of verbs are related to their semantics. She introduces the term 'alternation', under which she subsumes the possible changes in realization of the argument structure of verbs, sometimes accompanied by subtle changes of meaning. She defines 79 alternations. Verbs differ with respect to the number and types of alternations which they undergo. The verbs are grouped together into classes according to the combination of their possible alternations.

**Example:**

The class 'Verbs of Sending and Carrying' can be characterized by participation in:

dative alternation (which concerns verbs with two objects where the object attached by the preposition *to* can be expressed as the first object: *Bill sold a car to Tom. -> Bill sold Tom a car.)*

conative alternation (a subtype of transitivity alternations in which the object of a transitive construction is preceded by the preposition *at*: *Paula hit the fence. -> Paula hit at the fence.)*

causative alternation (a subtype of transitivity alternation, which has several subtypes and involves verbs which can be used transitively or intransitively; the intransitive usage is characterized by the emphasis on the cause: *Janet broke the cup.-> The cup broke.)*

middle alternation (a subtype of transitivity alternation that is created by transitive and intransitive usage of verbs; the intransitive usage is characterized by the absence of an agent and the presence of a modal or an adverbial element: *The butcher cuts the meat. -> The meat cuts easily.).*

The possibility or impossibility of changes in expressing arguments described by the above named alternations subdivides verbs of this class into more subtle classes:

'*Send* Verbs': *airmail, convey, deliver, dispatch, express,* etc. (dative alternation+, conative alternation-, causative alternation-, middle alternation-)

'*Slide* Verbs': *bounce, float, move, roll, slide* etc. (dative alternation+, conative alternation-, causative/inchoative alternation+, middle alternation+).

**2.2 PropBank[1],[2]** contains so-called PAS - predicate-argument structures. PropBank does not use semantically oriented labels and arguments of a given verb obtain a number from 0 to 5. A distinction between the obligatoriness and the optionality of arguments is not made. Each frame is represented by a list of possible complements of a given verb. The number of arguments is characteristic only for an individual verb or verb class.

**2.3 LCS[3],[4]** is based on semantically oriented considerations and works with a number of basic elements as types assigned (*Event*, *State*, *Path*, *Manner*, *Property* and *Things*), semantic fields (Locational, Possessional, Identificational, etc.) and primitives (structural primitives: CAUSE, GO, BE, TO and constants: reduce+ed, textile+, slash+ingly etc.). The main classification of verbs is based on Levin´s verb classes which are modified, i.e. furthermore subdivided into more subtle classes. Levin´s verb classes are assigned semantic structures, described by semantic structural primitives and semantic fields, whereas verbs differ from each other in their content specified by constants within a single class.

**2.4 FrameNet[5]** project is focused mainly on selecting words with particular meanings and describing the frames or conceptual structures underlying these meanings. In contrast to PropBank, the key role in FrameNet is played by semantics. FrameNet works with semantic frames, i.e. a schematic representation of situation types together with lists of the kinds of participants, and creates a hierarchy of semantic frames that are shared by groups of lexical units.

**Example:**

The group Cause-Motion: *cast, catapult, chuck, drag, draw, fling, hurl, nudge, etc.* is described by conceptual roles ('Agent', 'Theme', 'Source' / 'Path' / 'Goal') that are seen as components of such situations and are replaced by specific semantic arguments corresponding to the frame elements.

However inspirative these classifications are they seem to be hardly applicable to the material provided by VALLEX. This fact is caused by differences in the theoretical backgrounds and their methods and instruments of descriptions. Classification of verbs in VALLEX is based mainly on the syntactic criteria. However, we try to further exploit above mentioned methods of classifications. (Even though, we do not undertake their results.)

**A Brief Characterization of Functional Generative Description (FGD)**

FGD has been developed by Sgall and his collaborators since the sixties [Sgall, Hajičová, Panevová, 1986]. This framework of language description is based on a stratificational approach to the system of language. The highest layer is represented by the layer of the underlying structure that corresponds to the layer of linguistic meaning, the so-called tectogrammatical layer. Valency theory was designed as an important part of FGD and has been developed since the seventies, especially by Panevová [Panevová, 1974-75,1980].

The term 'valency' was used for the first time by Lucien Tesnière in connection with the dependency analysis of sentence [Tesnière, 1959]. By this term, an ability of verbs and other words, for instance nouns, adjectives and adverbs, to bind other words is generally understood. A valency frame is typically represented as a set of slots of syntactic elements (complementations, participants, modifications) required by a verb. On the tectogrammatical layer, five inner participants (*actants* in Tesnière's terms, *arguments* in more general terms) and free modifications are distinguished.

---

[1] http://www.cis.upenn.edu/~ace/

[2] http://www.cis.upenn.edu/~treebank/

[3] http://umiacs.umd.edu/~bonnie/LCS_Database_Documentation.html

[4] http://www.ilc.pi.cnr.it/EAGLES96/rep2node10.html

[5] http://icl.pku.edu.cn/doubtfire/semantics/FrameNet/Project/FrameDescs.htm

Inner participants satisfy the following two criteria:

(i) They cannot appear more than once within an occurrence of a particular verb (with the exception of apposition and coordination).

(ii) They can modify only some verbs. Therefore, the given combination of them is characteristic for a particular verb.

FGD distinguishes 5 inner participants: Actor/Bearer (ACT), Patient (PAT), Addressee (ADDR), Origin (ORIG) and Effect (EFF).

**Example:**
*Jana*.ACT *dostala od své matky*.ORIG *peníze*.PAT
Jane.ACT got the money.PAT from her mother.ORIG.
*Petr*.ACT *řekl Pavlovi*.ADDR, *že je hloupý*.EFF
Petr.ACT told Paul.ADDR that he was stupid.EFF.

On the contrary, free modifications can modify any verb and they can appear with a particular occurrence of the verb more than once. They correspond to adverbial complements and they are limited only by semantic restrictions. They can be subcategorized into several subclasses: complements of time, cause and consequence, manner, location and others. FGD distinguishes a relatively high number of types of free modifications (for the list of free modifications see e.g. [Mikulová et al., 2005]). In contrast to inner participants, most of free modifications are optional and their morphemic realization is not governed.

**Example:**
*Zítra*.TWHEN *ráno*.TWHEN *v deset hodin*.TWHEN *odjíždíme do Prahy*.DIR3.
We.ACT will leave for Prague.DIR3 tomorrow.TWHEN in the morning.TWHEN at 10 o'clock.TWHEN.

In FGD, the so-called shifting of inner participants is taken into account [Panevová, 1980]. FGD uses syntactic criteria for determining Actor and Patient whereas further inner participants (Addressee, Origin and Effect) are distinguished on the basis of semantic considerations.

Inner participants can be either (semantically) obligatory or optional and they are always a part of the valency frame of the given verb. Roughly speaking, obligatoriness of an inner participant means that it may not be omitted in the underlying structure. On the contrary, free modifications pertain to the valency frame only if they are obligatory. The so-called dialogue test is used as the criterion for obligatoriness of inner participants and free modifications [Panevová, 1980].

## 3. VALLEX, Valency lexicon of Czech Verbs and Verb Classes

The valency lexicon VALLEX[6] uses FGD as its theoretical background and it contains 2451 lemmas of Czech verbs and 6411 valency frames. VALLEX provides information on valency frames, synonyms, control, reflexivity and aspectual counterparts; also a distinction between primary, secondary or idiomatic usage is made (for more information see [Lopatková et al., 2002]). The valency frame is represented as a sequence of frame slots. Each slot represents one complementation and contains the functor, the list of possible morphemic realizations and the indication of the type of complementation (obligatory, optional, typical). Valency frames are assigned to individual lexical units that roughly correspond to the individual senses rather than to the lemmas.

Let us introduce a new trend in treatment of verbs in VALLEX, i.e. a tentative classification of verbs. At present, VALLEX contains 21 classes: **communication, mental action, perception, exchange, change, production, phase verb, phase of action, modal verb, motion, transport, location, expansion, combining, social interaction, providing, appoint verb, contact, emission, extent, psych verb.**

It is to be emphasized that the verb classes are built consistently from below. For the grouping of verbs, a number of mostly syntactically based criteria are taken into account. The main criteria for grouping verbs together are as follows: the combination of the **number of the complementations** (mainly obligatory, optional inner participants and obligatory free modifications), their type

---

(obligatory, optional, typical), **functors** (labels expressing the type of relation between verbs and their complementations) and their possible **morphemic realizations**. Further helpful criteria are reciprocity [Panevová, 1999], control [Panevová, 1996] and secondary diathesis.

VALLEX presents enriched valency frames that involve also the so-called typical modifications. Typical free modifications specify the whole groups of verbs. For instance, the class of verbs of **motion** is modified by typical free modifications of direction (Direction-from, Direction-through and Direction-to) that tend to occur together:

**Example:**
*Petr*.ACT *jel z domova*.DIR1 *přes celé město*.DIR2 *do obchodu*.DIR3
Petr.ACT went from home.DIR1 through the whole town.DIR2 to the supermarket.DIR3

We admit that at present the modifications are assigned on a rather intuitive basis. We expect that the classification of verbs can lead to an improvement of the data consistency in this task.

### 3.1 Classes with a Sentential Complement: Communication, Mental Action, Perception

Let us introduce some examples of groups of verbs with a sentential complementation: verbs of **communication, mental action and perception.** One of the complementations of the verbs in these classes can be realized as a finite clause.

**3.1.1** The group of verbs of **communication** involves a wide range of verbs, the so-called *verba dicendi*. Their frames may contain 3 or 4 participants. The first participant is the 'speaker', then there is an 'addressee' of the speech, a 'theme' and 'what is said about the theme'. Conjunctions introducing the dependent clause are listed in the slot representing the given type of complementation.

**Example:**
*Jan*.ACT *řekl Pavlovi*.ADDR *o své ženě*.PAT*, že ….*EFF
John.ACT told Paul.ADDR about his wife.PAT that … EFF;
*Jan*.ACT *řekl, že tam půjde*.EFF
John.ACT said that he would go there.EFF
ACT(1;obl) ADDR(3;opt) PAT(o+6;opt) EFF(4,aby,ať,že;obl)

**3.1.2** The group of verbs labeled as **mental action** is close to the verbs of communication - in their frames, one of the complementations may be expressed by a clause. However, the Addressee is missing within the valency frames of these verbs:

**Example:**
*Jan*.ACT *si uvědomil svou pozici*.PAT
John.ACT realized his position.PAT
*Jan*.ACT *si uvědomil, že je pozdě*.PAT
John.ACT realized that it was late.PAT
ACT(1;obl) PAT(4,zda,že;obl)

**3.1.3** In a similar vein, one of the complementations of verbs of the class **perception** can be realized as a clause and the Addressee is also missing as in the case of verbs of mental action. The difference is that the dependent clause can be introduced by 'jak' functioning as a conjunction rather than as a pronominal adverb, as is the case of the verbs of communication or mental action; therefore, with the verbs of perception, 'jak' is listed as a conjunction introducing the dependent clause, while with the verbs of communication and mental action, 'jak' functions within the dependent clause as one of the complementations of this clause. Consider the following examples:

*Petr*.ACT *Pavla*.PAT *viděl přicházet / jak přichází / že přichází*.EFF
Petr.ACT saw Paul.PAT come/as he comes/that he comes.EFF
ACT(1;obl) PAT(4;obl) EFF(inf,jak,že;obl)

as compared with:

*Jan*.ACT *řekl Pavlovi*.ADDR*, jak to má udělat*.EFF
John.ACT told Paul.ADDR how to do it.EFF
*Jan*.ACT *si uvědomil, jak to má udělat*.PAT
John.ACT realized how to do it.PAT

## 3.2 Further subclassification: Communication

We assume that some of these classes may be inherently heterogeneous and thus we expect that verbs may be subclassified into a hierarchy of classes that will correspond to the different levels of granularity. For instance, the large group of verbs of **communication** includes verbs with the meaning "presenting information to an addressee" and also with the meaning "giving an order". In contrast to the former group of verbs, the verbs of the latter group are characterized by the presence of the relation of control, i.e. one of their participants can be realized as an infinitive and the 'subject' of the infinitive is controlled by one of the complementations of the main verb. The controller of these verbs is the Addressee:

**Example:**
*farář*.ACT *jim*.ADDR *kázal jen samé nesmysly*.EFF
the parson.ACT preached them.ADDR only rubbish.EFF
ACT(1;obl) ADDR(3;opt) PAT(o+6;opt) EFF(4,že,rel;obl)
*Jan*.ACT *Pavlovi*.ADDR *kázal zahrnout*.PAT *ji dary*
John.ACT ordered Paul.ADDR to give.PAT her the gifts (i.e. who gives the gifts is Paul)
ACT(1;obl)  ADDR(3;obl)  PAT(inf,aby,ať;obl)

## 3.3 Differences in Morphemic Realization: Exchange

The morphemic forms of complementations can vary within the classes but the grouping of verbs can serve to list possible morphemic realizations of each complementation. Subsequently, checking whether an individual morphemic form is possible within a particular verb can be done in order to improve the data consistency.

Further examples represent verbs from the class of **exchange** that generally express a change of ownership. An ownership occupies the position of the Patient in both of the examples. However, we can distinguish several subclasses within the class exchange. For instance, the first two examples describe a change when the ownership passes over from the Addressee to the Actor (1) whereas the Addressee acquires the ownership from the Actor in the second couple of the examples (2). Thus, these couples are characterized by the same semantics but they differ in morphemic forms of participants ((1) ACT-Nominative, ADDR-Dative (i) / Accusative (ii), PAT-Accusative (i) / o+Accusative (ii), (2) ACT-Nominative, ADDR-Dative (i) / Accusative (ii), PAT-Accusative (i) / Instrumental (ii)). Thus, morphemic realization of inner participants can vary within these two 'subclasses'[7]:

**Example:**
(1) (i) *Jan*.ACT *vzal Pavlovi*.ADDR *peníze*.PAT
    John.ACT took Paul.ADDR the money.PAT
    ACT(1;obl) ADDR(3;obl) PAT(4;obl)
  (ii) *zloděj*.ACT *okradl Jana*.ADDR *o peníze*.PAT
    the thief.ACT robbed John-Accusative.ADDR of the money-Accusative.PAT
    ACT(1;obl) ADDR(4;obl) PAT(o+4;obl)
(2) (i) *Jan*.ACT *dal Marii*.ADDR *dárek*.PAT
    John.ACT gave Mary-Dative.ADDR a gift-Accusative.PAT
    ACT(1;obl) ADDR(3;obl) PAT(4;obl)
  (ii) *Pavel*.ACT *obdaroval Marii*.ADDR *dary*.PAT
    Paul.ACT gifted Mary-Accusative.ADDR with the gifts-Instrumental.PAT
    ACT(1;obl) ADDR(4;obl) PAT(7;obl)

## 3.4 A Borderline Case between Inner Participants and Free Modifications: Verbs of Providing

Classifying verbs also opens the question of a borderline between free modifications and inner participants. The main criteria for distinguishing inner participants and free modification [see Section 2, for further information see Panevová, 1980] need refinement in some cases [see Urešová]. For instance, the class of verbs of **providing** covers verbs with the meaning "providing something with

---

[7] For further details, see [Benešová, 2004]

something (else)". The third complementation can be classified either as an optional participant or as a typical free modification. Classifying verbs allows to solve the similar questionable examples in a consistent way.

**Example:**
*naplnil láhev*.PAT *mlékem*.EFF/MEANS
ACT(1;obl) PAT(4;obl) MEANS(7;typ) / EFF(7;opt)
He.ACT filled a bottle.PAT with milk.EFF/MEANS

## 4. Conclusion

We have attempted to formulate basic criteria for the classification of verbs (according to the combination of the number of their complementations, of their types, functors and their possible morphemic realizations) and to present further helpful criteria for such a classification (e.g. control). We have also paid an attention to the heterogeneity of the classes. We have demonstrated on the verbs from the same (sub)class that the morphemic realizations are not the primary criterion for assigning their semantic roles (functors in FGD). We have also mentioned some open questions (the borderline between inner participants and free complementations) that have appeared during the classification of verbs. In conclusion, as far as future plans are concerned, we would like to step forward in building a hierarchy of verb classes with regard to their similarity ('more general classes') or to the dissimilarity of verbs within a single class ('more subtle classes') and in checking the consistency of data on the basis of these classes.

## References

Benešová, V., Delimitace lexií českých sloves z hlediska jejich valenčích vlastností. diploma work, Charles Univerzity, Faculty of Philosophy and Arts, Institute of Czech Language and Theory of Communication, Prague, 2004.

Levin, B., English Verb Classes and Alternations: A preliminary Investigation, *Chicago, University of Chicago Press*, 1993.

Lopatková, M., Z. Žabokrtský, K. Skwarska, V. Benešová, Tektogramaticky anotovaný valenční slovník českých sloves, ÚFAL/CKL Technical report TR-2002-15, 2002.

Mikulová, M., et al., Anotace na tektogramatické rovině Pražského závislostního korpusu, 2005.

Panevová, J., More remarks on control, Prague Linguistic Circle Paper 2, John Benjamins, Amsterdam, 101-120, 1996.

Panevová, J., On Verbal Frames in Functional Generative Description, Part I, *PBML*, 22, 3-40, Part II, *PBML*, 23,17-52, 1974-75.

Panevová, J., Formy a funkce ve stavbě české věty, *Praha, Academia*, 1980.

Panevová, J., Česká reciproční zájmena a slovesná valence, *SaS*, 4, 1999.

Sgall, P., E. Hajičová, J. Panevová, The Meaning of the Sentence in Its Semantic and Pragmatic Aspects, D. Reidel Publishing Company, Dordrecht, 1986.

Tesnière, L., Éléments de syntaxe structurale, Paris, 1959.

Urešová, Z., The verbal valency in the Prague Dependency Treebank from the anotator's point of view, Bratislava, in print.