# The communicative potential of verbs of "away-from" motion in English, German and Russian*

Serge Sharoff
University of Leeds

This paper proposes a model for describing lexical semantics within systemic-functional theory. Formal approaches to lexical semantics assume that words have meanings which exist independently from communication and are represented by means of (semi-) formal definitions. Functional approaches treat language as a tool for social interaction and, in the case of lexical semantics, assume that lexical items provide resources for realising the communicative intentions of the speaker. The paper explores problems occurring when uses of words in context are categorised according to pre-established sense distinctions and proposes a mechanism that uses systemic networks to describe meanings of lexical items as functions of their uses. The paper ends with a case study of uses of English, German and Russian verbs of motion, including options for their translation. The model is also tested empirically by considering word uses in an aligned parallel English-German-Russian corpus.

## 1. Introduction

Most contributions of systemic-functional linguistics (SFL) in the domain of lexicogrammar so far have concerned the analysis of grammatical structures; i.e. relatively few efforts within the SFL paradigm have been devoted to studying lexical meanings. Yet the distinction between the lexicon and the grammar was clear from the onset of systemic descriptions. According to Halliday (1961), word is a **grammatical** category within the ranks of clause, group, word, morpheme. A word is defined in terms of its contribution to the meaning of a group and integrates contributions from morphemes. In contrast, lexical item is a **lexical** category and can appear at any rank of the grammatical description.

*Functions of Language* 12:2 (2005), **203–238**.
ISSN 0929–998X / E-ISSN 1569–9765 

For example, the following expressions can be considered as lexical items: *a dialogue box* (a nominal group), *How are you?* (clause), *-able* (morpheme).[1] When a lexical item, e.g. a nominal group, is embedded in a grammatical environment, its elements carry syntactic functions, for example, *dialogue* is Classifier in *the dialogue box* (Halliday, 1985:146). In English, the lexicogrammatical realisation of *the dialogue box* is not significantly different from that of *the box*. The lexical item *dialogue box* can even be represented in the grammatical structure as a single unit (a word with a space inside). However, in Russian the modifying adjective agrees in case and number with the noun head, e.g. *dialogov<u>oe</u> okn<u>o</u>, dialogov<u>ogo</u> okn<u>a</u>, dialogov<u>ym</u> okn<u>om</u>*. Thus, both grammatical structures and lexical meanings should be represented at different, but intersecting levels of description.

However, this exposition leaves open the question of the nature of lexical studies within SFL. The answer is hinted at by the opposition of logical and ethnographical-rhetorical paradigms in linguistic research (Matthiessen and Bateman, 1991:54–55). In the study of grammar, the two perspectives are partly mirrored in the opposition of formal and functional linguistics. In the study of lexis, the first paradigm assumes that meanings are **concepts**, which represent real-world objects and their properties. Among computational approaches to lexical semantics, good representatives of this paradigm are Meaning-Text Theory, MTT, (Mel'chuk, 1988) and WordNet (Miller, 1990). Even though the spectrum of approaches within this paradigm is wide, and some theories are in opposition to others, all approaches that can be considered as belonging to it share the same ontological commitment: the relationship between words and respective senses is primary with respect to communication, since the ontology, as a repository of senses, exists and the mapping from words to senses is defined before any possible act of communication. In contrast, the second paradigm assumes the primacy of communication, because it holds that human languages are not informal prototypes for knowledge representation languages, but tools that enable both the communication of experience and interaction in society. Thus, the meanings of lexical items are analysed in terms of their contribution to the exchange between the speaker and the hearer.

The first paradigm assumes that a word has a meaning before any possible use, while the second one treats the meaning of a word as a **function** of its use. There is also a terminological difference between the two approaches with respect to the topic of study. Logical approaches prefer to use the term 'sense', which invokes the dictionary metaphor: a word refers to a set of senses in a dictionary. In functional approaches, the term 'meaning' is more appropriate,

which follows the metaphor of communication: when a word is used, it carries a meaning intended by the speaker. Throughout the present paper the terms 'sense' and 'meaning' are retained for formal and functional approaches respectively.

One comprehensive account of the English lexicon, which is functional and cognate to systemic linguistics, is the Collins Cobuild English Dictionary (CCED) and the pattern grammar approach (Hunston and Francis, 1999) which followed from it. However, descriptions in a human-oriented dictionary lack formal mechanisms for dealing with meanings of lexical items in a computationally tractable way, i.e. such descriptions cannot be directly used in computational applications for language understanding and generation or machine translation. The model proposed in this paper is based on the systemic choice network that allows a formal computational implementation.

In Section 2, I consider the "meaning-as-concept" model of lexical items and the aspects of language use that are not covered by it. Then, Section 3 presents a framework for describing meanings of lexical items as uses by means of systemic networks. Finally, Section 4 reports a comparative study of uses of English, German and Russian verbs of motion, in particular options referring to physical and abstract motion, as well as an empirical study of translation equivalents in an aligned parallel corpus.

## 2. Words as pointers to senses

According to the logic-centered paradigm, the relationship between words and meanings is a one-to-many mapping between a lexical item and a set of its senses. Each concept in the list of senses is a separate item with necessary and sufficient conditions for considering a use as an instance of the concept. The difference between approaches in this paradigm concerns their understanding of what constitutes a sense.

WordNet uses the idea of the lexical matrix, in which senses are encoded by synsets, i.e. sets of synonymous lexical items (Miller, 1990). Figure 1 demonstrates the principle of the lexical matrix: all senses that exist in English are listed as $M_1 \ldots M_m$; all English words are listed as $F_1 \ldots F_n$. If there are two entries in the same column of the matrix (e.g. $E_{1,2}$ and $E_{2,2}$ in Figure 1), the word form ($F_2$) is polysemous; if there are two entries in the same row (e.g. $E_{1,1}$ and $E_{1,2}$), the two words ($F_1$ and $F_2$) are synonyms with respect to this sense. For each lexical item, WordNet lists its synsets and often gives them a definition in

| Word Meanings | Word Forms | | | | |
|---|---|---|---|---|---|
| | $F_1$ | $F_2$ | $F_3$ | … | $F_n$ |
| $M_1$ | $E_{1,1}$ | $E_{1,2}$ | | | |
| $M_2$ | | $E_{2,2}$ | | | |
| $M_3$ | | | $E_{3,3}$ | | |
| ⋮ | | | | ⋱ | |
| $M_m$ | | | | | $E_{m,n}$ |

**Figure 1.** The Lexical Matrix in WordNet

informal English; however, a synset in WordNet is not defined, but signified by the set of its lexical items. For instance, one synset for *leave* in WordNet 1.6 is: "leave, go forth, go away (go away from a place)".

MTT provides a specific lexicographical device for representing senses of lexical items, the Explanatory Combinatorial Dictionary (ECD). An entry for a separate sense consists of the definition zone, in which the sense is defined in terms of semantically simpler components, which are either primitive or have their own definitions (they should be ultimately reduced to primitives). Mel'chuk (1988:173) defines the following principle: "In the definition of a lexical unit L, each component must be necessary, and the set of all components must be sufficient, for the definition to identify L uniquely in all imaginable uses." The definition is expressed in a natural language (English, French or Russian), but the ultimate goal for representing senses in ECD is to use a formal knowledge representation mechanism, like predicate calculus or semantic networks, e.g. each sense can be described by a conceptual graph (Sowa, 1991). However, when real knowledge-representation languages are used, the set of senses is typically reduced: a word refers not to 17 senses of *leave* in WordNet or 29 senses in CCED, but to a single predicate or a piece of the semantic network, for instance, Miller (1972) defines *leave* as simply (away(travel))($x$).

The ontological commitment, namely that senses of lexical items are concepts taken from an ontology, is based on the dictionary metaphor: using words is similar to using dictionaries. The set of senses of a lexical item in a monolingual dictionary is an informal prototype for the description of cognitive operations performed by humans. However, the concept and the format of modern monolingual dictionaries depend on the history of their development and on the function they serve in society. As Kilgarriff (1997) points out, origins of the tradition of sense enumeration in a monolingual dictionary are related to the development of printed discourse, particularly new periodicals, and the growth of a literate population in England at the beginning of the eighteenth century.

This brought about a reevaluation of the nature of meanings: dictionaries appeared as a response to the demand for an authoritative source which helps in understanding the use of a word or in checking/ensuring the correctness of its use. As the result of this intention, a dictionary entry is typically designed as a list of senses that denote objects or their properties.

This design is suitable for many purposes, but it can lead to descriptive problems when the task is to develop a model of uses of words in discourse. The following subsections illustrate two types of problems: how to decide which sense is to be selected for a given use (Section 2.1) and how to match different sets of senses of lexical items in translations (Section 2.2).

### 2.1 Mismatches between concepts and uses

According to the logic-centered paradigm, when a word is used in an utterance, it refers to one concept that is selected from the list of its senses. The mapping is based on the assumption that all possible uses of a lexical item should be expressed explicitly in its definition, which lists necessary and sufficient conditions for it referring to each and every concept. The drawback of the model is the possibility that the list of senses can be simultaneously too short and too long, once real uses are considered. It can be too short, because there are typically more possibilities for using a word than defined by a concept. This topic has been widely discussed in linguistic research: it is an immensely difficult task to construct a "waterproof" definition with a set of necessary and sufficient conditions for considering specific senses. A carefully constructed definition often fails when confronted with possible uses, cf. the discussion of *bachelor* in Winograd (1976) or examples with *leave* and *vyjti* below. At the same time, if the list of senses gets more elaborate to accommodate all possible uses, it starts to be too long, since some uses may appear to be ambiguous, i.e. referring to more than one sense in the list, while they are not ambiguous for either the speaker or hearer.

As an illustration of these problems, I analyse definitions of two close translation equivalents: *leave* in English and *vyjti* in Russian. The definitions were taken from WordNet 1.6 (see Appendix 1) and Apresjan (2000) (see Appendix 2). As is typical with a lexicographical treatment, there are errors or omissions that allow for a simple correction. For instance, Apresjan considers the following sense of *vyjti*: 6. *vyjti zamuzh*, to get married [of a woman]. This use is highly idiomatic and in the proper ECD style should be expressed by a separate lexical entry. At the same time, Apresjan (2000) lists the following uses

as idiomatic and does not include them in the list of senses: *vyjti na pensiju* (to retire) and *vyjti v otstavku* (to resign from one's work). The list can be easily extended, for instance: *vyjti v rezerv* (to be transferred to the reserve temporarily), *vyjti v zapas* (to be transferred to the reserve). In the ECD style, they could correspond to a separate sense "to leave an institution".

However, other problems do not allow for an easy solution. There is no way to classify examples (1–3) in terms of the senses from Apresjan (2000):

(1) … *ne* ***vyhodja*** *iz programmy.*
… not exiting from the application.
(2) *Pri vhode i* ***vyhode*** *iz polja formy …*
On entering and exiting a form field …
(3) *Doroga* ***vyhodit*** *k reke Sveltaja i idet vdol' nejo 10 km.*
the road reaches (lit. "goes out") to river Svetlaja and follows it for 10 km.

The most natural is to apply the sense 1.2 "to move outside", but the complete definition explicitly states: "a vehicle moves into a more open space". Example (1) does not refer to vehicles, but still may be interpreted as a transition from a closed space (an application) into an open space (the operating system). However, even this metaphorical reinterpretation is impossible in (2), since the transition is made between two fields, and in (3), where the road (not a vehicle) follows a path and does not reach a more open space. Anyway, the need to use metaphoric interpretations in these cases violates the requirement of explicit and rigorous definitions. In addition, VYJTI 1.2 states that the agent is necessarily a vehicle, but an example of this given in Apresjan (2000) is *Karavan vyshel iz oazisa* (the caravan left the oasis), which does not refer to a vehicle and, most probably, not even to a group of vehicles according to its possible ECD definition.

Similar problems occur with more abstract senses of *vyjti*. For instance, the following use of *vyjti* from a Russian translation of Lewis Carroll's 'Alice in Wonderland':

(4) *iz nego by* ***vyshel*** *ves'ma neprijatnyj rebenok.*
from him would go-out rather ugly child-NOM
'it would have made a dreadfully ugly child'

initially appears to refer to senses 7.1 or 7.2, but it violates the condition of increased social status in 7.1 and gaining expertise in 7.2.

On the other hand, many examples refer to "in-between" cases, when a use corresponds simultaneously to several senses (according to WordNet for *leave* and Apresjan, 2000, for *vyjti*). For instance, sentence (5):

(5) *te s radost'ju zapisali vse tri otveta, potom slozhili ih i **vyshlo**: 44 kopejki.*
'they (the jury) eagerly wrote down all three dates, and then added them up, and the result came: 44 kopecks.'

may refer to sense 7.3 (production of something) and sense 7.4 (acquisition of a property), though neither sense fits completely: the definition of 7.3 assumes the production of an object from some raw material, while the definition of 7.4 assumes the possibility for an object to acquire one of several properties by chance. *Vyjti iz boja* (to leave the battle field) is used as an example in 3.3 (to cease to do), but it also corresponds to physical motion in 1.1. Another example of 3.3 from the military domain, *vyjti iz okruzhenija* (to extricate oneself from the encirclement) is related in addition to physical motion to sense 3.2, i.e. to cease to be in the state of encirclement.

Apresjan (2000) offers a thorough treatment of a single word, so we can expect similar problems in a wide-coverage dictionary, which pays much less attention to individual words. In particular, WordNet contains two subsets of senses of *leave*: "go away from a place" and "move out of, as of a room, a country, a bus, etc." In many utterances the distinction between them cannot be drawn, so the use can only be considered as ambiguous. Also, Sense 9 (leave alone or undisturbed) is claimed to be synonymous with Sense 4 (leave unchanged or unaltered). The relationship between the two senses is indeed evident, but the nature of synonymy between synsets is unclear. Another group of almost synonymous synsets is Sense 2 (leave behind), Sense 11 (refrain from taking) and Sense 16 (leave behind, forget). They often refer to the same situation: an object remains after a person has left. The only difference is the emphasis on the deliberate character of the action in $leave_{11}$ and its inadvertence in $leave_{16}$, while $leave_2$ is their common denominator. In CCED, the three senses are combined into one description ($leave_4$).

Finally, even non-synonymous senses can be used quasi-ambiguously, i.e. can be applicable to a single use, as in (6):

(6) *The clues he left did not establish his identity but they reflected his personality*

this refers to Sense 2 (leave behind) and Sense 7 (*The water left a mark on the silk dress*). On the other hand, WordNet (unlike many other dictionaries) does not distinguish between several apparently different uses of *leave* in Sense 2, which are summarised in the definitions from CCED as:

(7) *If you **leave** a place or person, you go away from that place or person.* vs.

(8) *If you **leave** your husband, wife, or some other person with whom you have had a close relationship, you stop living with them or you finish the relationship.*

WordNet 1.6 has a semantic concordance consisting of some texts from the Brown Corpus manually annotated with WordNet senses. Even though the level of ambiguity reported in the WordNet concordance is low (703 multiple senses for 234136 tagged word forms), a closer look at the data shows that the real level of ambiguity is significantly higher. For instance, the distinction between $leave_2$ (leave behind), $leave_3$ (cause to be in a specified state) and $leave_4$ (leave unchanged or unaltered) cannot be drawn in many cases. An example for $leave_2$ in WordNet is *to leave a mess,* while for $leave_3$ is *to leave money.* The following example from the WordNet semantic concordance:

(9) *The general … explanation has left basic positions unchanged*

is tagged as $leave_3$ (cause to be in a specified state), but $leave_4$ is also applicable (leave unchanged or unaltered).

A concept in the list of senses may be related to other concepts, which are other senses of the same word, but the relationship is rarely specified explicitly in its definition. Senses in WordNet are designed as separate synsets. The ECD introduces the apparatus of semantic bridges in definitions, i.e. two related senses of a lexical item must share a non-trivial semantic component. However, some bridges are introduced artificially, i.e. a component is introduced for the sole purpose of not splitting the entry into two homonymous lexical items. For instance, Mackenzie and Mel'chuk (1988) add the component "X is very full of Y" for PREGNANT.1 (X is a woman or female animal, Y is a fetus) in order to share it with PREGNANT.2 (information object X is characterised by, contains or entails much Y/many Ys, as if X were very full of Y). Firstly, the bridge component is not always relevant for PREGNANT.1, for example, the joke Mackenzie and Mel'chuk (1988) use as their starting point implies that a woman is pregnant for only ten minutes. Secondly, the bridge component is not intended in many uses of PREGNANT.2. For example, the British National Corpus (BNC) lists (10)–(12):

(10) *It is a dangerous moment, pregnant with hope teetering on the edge of despair.*
(11) *It's a grey day, pregnant with rain.*
(12) *Oh which is the one that is pregnant with the terrible time?*

Incidentally, none of the sentences (10)–(12) refers to information objects, thus violating necessary conditions in the definition of PREGNANT.2.

A step towards a more dynamic account of word meanings is offered by the Generative Lexicon (GL) theory (Pustejovsky, 1995). According to GL, polysemous senses are represented by combining regular uses into an underspecified type, which is made more specific in the context. For instance, the word *book* may be interpreted either as a physical object or as a piece of information, because the function of a book is to be read by a person (this is represented by the Telic role) and it is produced by another person (this is represented by the Agentive role). So the type *book* can be defined as follows:

(13)
$$\begin{bmatrix} book(x) \\ formal = phys\_obj(x) \\ telic = read(P,y,x) \\ agentive = wrote(T,w,x) \\ constitutive = bind(x,z) \end{bmatrix}$$

However, even though Pustejovsky's theory is aimed at describing the dynamic account of sense production instead of sense enumeration (thus, it considers how a word is used), it still belongs to the logical side of the continuum, since meanings of a word in GL are considered from the viewpoint of their logical structure, not from the viewpoint of their use in communication and possible intentions of the speaker. Indeed Pustejovsky's proposal is similar to the notion of regular polysemy introduced within the MTT framework, (cf. Apresjan, 1974). At the same time, the particular set of senses of the noun *book* is not determined solely by the logical structure of the concept BOOK. For example, the sense $book_5$ in WordNet is "a record of commercial accounts", which is treated as a separate sense, but its GL definition is indistinguishable from the definition for $book_1$ (printed work) cited in (13), because $book_5$ is also a physical object, produced by writing and aimed at reading (though of a different sort), so (13) joins two apparently different concepts. Admittedly, $book_5$ is always used in the plural and other dictionaries do add a restriction (*pl.*), but it is harder to express this fact in the more formal structure of WordNet, which is designed as a mapping between the list of words (defined as lemmas) and the list of senses (cf. Figure 1). In addition, $book_8$ from WordNet (a book of tickets or stamps) is not produced by writing and is not aimed at reading, and even though it is constituted by pages bound together like $book_2$, the notion of pages, the way they are bound and their function in $book_8$ is different. WordNet does not specify that the word in this sense is always used in plural (*books*).

### 2.2 Mismatches between concepts in translation

Reference to the set of senses should also be preserved when a lexical item is translated into another language. This implies that the inventory of concepts is essentially shared across languages, which use different words for pointing to them. However, the task of meaning preservation in translation is not trivial. For instance, the Collins English-German dictionary distinguishes 19 senses for the translation of *leave* into German and lists 29 different words which can be used as translation equivalents. Some differences between translation equivalents are related to the inherent ambiguity of *leave* in English: it does not specify the means of leaving, while this is often obligatory in German and Russian. For instance, *When he left Rome* should be disambiguated in German and Russian translations (14):

The German translations of *left home* also depend on whether this means a permanent departure (*weggehen*) or a short-term departure (*wegfahren*). Thus, the need to translate a word into a target language multiplies the number of its senses.

(14) *Als er von Rom wegging / wegfuhr / abflog …*
*Kogda on iz Rima ushel / uehal / uletel …*
when he from Rome went-away drove-away flew-away

Finally, even if a use is not ambiguous, it can refer to slightly different concepts in different languages. For instance, actions with buttons and choice lists as control elements of software applications are typically expressed in English by using the verb *choose* as a logical choice among available options:

(15) *Choose the Save button*
(16) *Choose File:Save*

However, in Russian, choosing a semi-real button in a dialog box is expressed as a physical action of pressing:

(17) *Nazhmite knopku Save*
press-IMPER button-ACC Save
'Press the Save button'

but choosing an item in a menu is expressed as a logical choice:

(18) *Vyberite punkt File:Save*
select-IMPER item-ACC File:Save
'Select the File:Save item'

*Click* is a more specific verb for the action referred to in (15), but it assumes that the action is performed using a mouse, while *choose* allows the possibility of using a keyboard for doing this. Even though, the verb *vybirat'* in Russian is the typical translation equivalent of the verb *choose* in many contexts, it is not appropriate in (17), most probably because it does not imply an act of option selection together with its functional consequences, as in *choose*. Thus, the senses of the verbs *choose* and *vybirat'* in (15) and (18) should be represented by two different concepts. Such minor differences between senses of similar words in different languages lead to the separation of sets of their senses.

## 3. Words as resources for communication

The conclusion from the description presented in the previous sections is mostly negative. It confirms that there is a mismatch between concepts defined as senses in a dictionary and uses of words treated as instances of such concepts. What are the implications of these problems for modeling lexical meanings in the framework of systemic linguistics? After all, any dictionary is finite (in practice) in comparison to an infinite amount of situations in which words can be used. How can we describe lexical items in terms of their uses, in particular in tasks oriented to natural language processing?

The essence of the proposed approach to describing meanings of lexical items from a communication-centred perspective is presented in Figure 2. It assumes two parties involved in communication: the speaker with communicative intentions and the hearer. The speaker uses lexicogrammatical resources

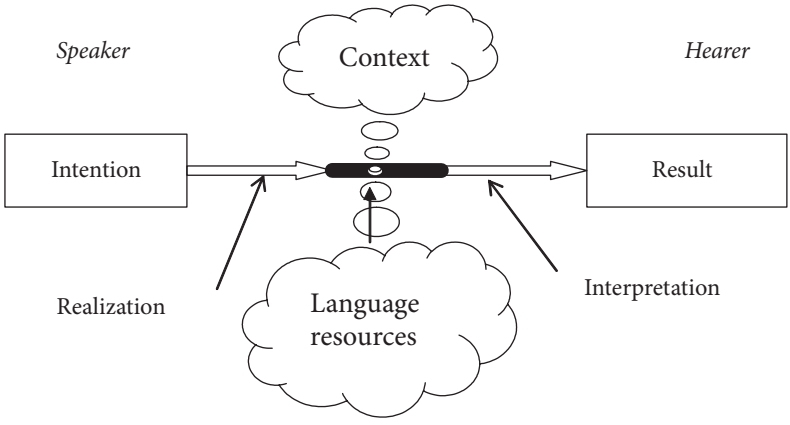


**Figure 2.** The communication-centered approach to uses of lexical items

for expressing communicative intentions according to the context of situation. The meaning intentions are realised as sequences of physical signals (printed characters, acoustic waves, non-verbal gestures), which are interpreted by the hearer. Interpretation is based on the set of linguistic resources and the context of situation, which in the case of successful communication are largely shared with the speaker. The description of meanings of lexical items in terms of their uses consists in defining the **potential** of possibilities for realising meaning intentions by means of lexical items and in describing the **instantiation** of the potential, when lexical items are used in context. Thus, the approach is, in a sense, generative: the finite set of lexicogrammatical resources is used to communicate a large set of speaker's meaning intentions in an infinite amount of contexts.

Halliday and Matthiessen (1999:504) offer the following trinocular view:

1. "from around" the lexicogrammar: to consider what choices are available in the lexicogrammar, for instance, [declarative] and [interrogative] within the system of indicative clauses;
2. "from above" the lexicogrammar: to look at the distinction in meaning which is reflected in the choice between [declarative] and [interrogative] in the system;
3. "from below" the lexicogrammar: to consider differences between [declarative] and [interrogative] in terms of their syntagmatic realisation.

In the case of lexical semantics, the view "from above" means a description based on communicative intentions;[2] the view "from below" is based on individual lexical items, while the view "from around" describes choices available for uses of lexical items in a lexical domain (verbs of motion, size adjectives, words denoting emotions, etc).

The view "from above" the lexicogrammar is based on information available in the context of situation. In the case of a computational model, sources of information can be implemented by various semantic inquiries, which relate options available in the systemic network to properties of the situation (Matthiessen and Bateman, 1991). For instance, such principles as the presence or absence of the speaker and/or the hearer at the destination point can guide the choice of *come* vs. *go* (Fillmore, 1983), or the difference in the type of the carrier of motion, e.g. motion by foot, wheels, wings, boat, etc, can guide the choice of *go, drive, fly, sail*. However, the choice is not straightforward, it also depends on traditions for expressing certain events and objects and the appropriateness of an expression in the current discourse context. For instance,

the domain of flight, bus or ship schedules in English favours generic terms for designating the departure of a vehicle, like *leave* (*my flight leaves in less than an hour*). In German and Russian, the possibilities for using generic references are more restricted. In Russian the departure of a plane should be referred to as flying (*vyletat'*), while other types of departure are expressed in a generic way (*vyhodit',* leave). In German all cases of departure according to a schedule are expressed by verbs specific to the type of a vehicle: *abfahren* (leave by bus, train and boat)[3] and *abfliegen* (leave by plane).

The view "from around" the lexicogrammar is based on the systemic network, which includes **paradigmatic** classifications of features and **syntagmatic** realisations, which specify consequences in the choice of features. In the case of lexical semantics, realisation statements constrain the choice of lexical items from the lexical stock available in a language. The potential of uses is represented by choices of features in the systemic network (feature names are shown in square brackets in the text).

According to the same principles as those used for the development of systemic grammars (Martin, 1987), a feature in the lexical network is motivated if it has some reflex in form. Thus, if several lexical items are available for expressing a group of communicative intentions, they correspond to several distinct features, e.g. *go, drive, fly* for distinguishing between vehicles. Equally, there may be grammatical constraints on realisation following the choice of a feature, e.g. the source or the destination is obligatory. In principle, each type of use should correspond to a unique set of features, which should be related to respective communicative intentions. For instance, *escape, move out, quit, run away* can refer to leaving a place without an intention to return. However, the verbs provide different lexical realisations for different communicative intentions, respectively, leaving a place of confinement (*escape*), relocating to a new home (*move out*), leaving a place under external pressure (*quit*), escaping from home secretly (*run away*). Such properties of the place or the cause of leaving are reflected in features of the systemic network, which result in the selection of the lexical item best suited for expressing the intention.

The lexical network also addresses the issue of differences between uses of words in different languages. Systemic research on the grammar of typologically related and unrelated languages shows that less delicate, more general, choices tend to be shared across languages, while more delicate choices tend to be language-specific (Bateman *et al,* 2000). The same principles guide the description of lexical meanings by means of the systemic network. For instance, many languages have the type of 'motion verb' and also distinguish between

motion towards or away from a reference object, e.g. *enter* vs. *leave, eingehen* vs. *ausgehen, vojti* vs. *vyjti*. Moreover, the three languages in this study have specific options to express the fact that a person leaves a vehicle at the end of the journey, e.g. *get out*, not *go out* or *leave*, but further options are language-specific: English distinguishes between public and private transportation means (*get off a train, bus, plane* vs. *get out of a car or boat*), German uses *aussteigen* for all types of vehicles, while Russian distinguishes between the relative height of the vehicle: *vyjti* for any vehicle, but *soyjti* for leaving a vehicle that is higher than the ground (mostly planes and ships, but also, less frequently, trains). Thus, language-specific systems and features should be added to the lexical network, which then represents *both* commonalities *and* differences between languages.

Other scholars have used systemic networks as the formal mechanism for describing lexical items. This includes the study of lexis as most delicate grammar (Hasan, 1987), the study of lexical options for adjectives within the lexicogrammatical network (Tucker, 1998) and the polysystemic model, treating lexis as a separate stratum (Wanner, 1997). These approaches address some properties of the behaviour of lexical items, but remain restricted in certain respects. Hasan (1987) treats few lexical items, like *strew, spill,* etc, which are also quite rarely used. Tucker's model covers the whole domain of adjectives, but it suffers from cursory treatment of the complete range of their uses. For instance, in the case of size adjectives it does not address such cases as *a broad coalition, a little girl,* or *high interest rates*. Wanner's model uses a separate stratum of lexis, which provides resources for converting situation specifications into lexicalised semantic specifications consumed by the grammar.[4] The separation of the lexicon and the grammar can lead to reduplication of grammar-induced properties in the lexical network and vice versa. Also Wanner (1997) describes general principles of his model, but does not provide a detailed study of a set of lexical items in terms of the model. There is also one criticism applicable to all these three approaches: unlike CCED, they are not functional enough, in the sense that they do not relate lexical items to communicative functions underlying their uses. The study of verbs of motion reported below is an attempt to overcome these limitations.

## 4. Towards the comparative lexical cartography of verbs of motion

This section presents a case study in the domain of verbs expressing "away from" motion in English, German and Russian. The selection is based on the aim of covering most frequent uses of most frequent verbs of motion according to representative corpora. Thus, the study is designed to mix considerations from form and function: as with any corpus study, it is based on (lemmatised) word forms found in corpora, but the aim of the study is to describe meaning intentions underlying uses of word forms. For instance, the consideration of which verb of motion is frequent is based on word frequency lists from the respective corpus: the British National Corpus for English, the XLEX corpus for German and the Russian Reference Corpus for Russian. Some information on types of their uses can be taken as senses from dictionaries: CCED for English, Wahrig for German and Ozhegov for Russian. However, very little information on the frequency of senses is provided in dictionaries. CCED makes an attempt to order senses according to their frequency, but it does not pursue this consistently (other principles intervene, e.g. physical senses often go first), and it does not give any frequency estimations for the senses. Moreover no reliable statistics are available even on the frequency of word forms for phrasal verbs in English and verbs with separable prefixes in German (particles and prefixes frequently code the direction of motion).

Table 1 lists all verbs used in the study. To achieve wide coverage of the domain, the study used all English, German and Russian verbs with frequency above 25 instances per million words (this corresponds approximately to the 3000 most frequent words in a language) and their combinations with most frequent particles and prefixes referring to "away from" motion. Initial types

**Table 1.** The most frequent verbs of "away-from" motion in English, German and Russian

| | |
|---|---|
| English | *come off, come out, drive away, drive off, drive out, escape, fly away, fly off, fly out, get away, get off, get out, go, go away, go out, leave, move away, move off, move out, pull out, quit, run away, run off, run out, walk away, walk off, walk out, withdraw* |
| German | *abfahren, abgehen, abkommen, ablaufen, ausfahren, ausgehen, auskommen, auslaufen, aussteigen, austreten, gehen, sich entfernen, entgehen, entkommen, herausgehen, laufen, verlassen, verkommen, verlaufen, wegfahren, weggehen, weglaufen, wegkommen, wegziehen* |
| Russian | *ishodit', otojti, otpravit'sja, poehat', pojti, pokinut', poletet', sbezhat', sojti, ubezhat', udalit'sja, uehat', ujti, uletet', vybrat'sja, vyehat', vyjti, vyletet', vylezti* |

of uses were provided by dictionary senses, which were extended by attempts to code examples of uses in corpora. Translations of verbs under consideration were also studied on the empirical basis of a corpus of aligned parallel texts and multilingual concordancing software (Sharoff, 2002). This corpus consists of several technical texts with descriptions of software, e.g. Microsoft Word Help, and literary texts, e.g. *Alice's Adventures in Wonderland* by Lewis Carroll. For the latter text, a German translation was also included (in addition to its three Russian translations).

The lexical semantics of verbs of motion is a popular research topic (see, for instance, Miller, 1972; Talmy, 1972; Levin and Hovav, 1992, for English; Eichinger, 1989; Gerling and Orthen, 1979, for German; and Apresjan, 2000; Grenoble, 1986, for Russian). Examples with such verbs are also widely used in general discussions on lexical semantics, for example Bierwisch and Schreuder (1992) and Langacker (1988). However, most analyses miss the complete range of options which are available for using verbs of motion; in particular, they pay little attention to processes which are typically expressed as a motion, but do not refer to a change in physical location, for example, *to go through a lot of trouble, to advance the claim, to pull out of the project*. Even for physical processes, many studies deal with basically one meaning for a verb and miss other possible uses. For instance, Miller (1972) considers a class of traveling verbs and defines *to go* simply as "to travel away from the speaker", thus, missing many other types of uses, like *It took us an hour to go three miles* or *I went home at the weekend*. In contrast, the present study shifts the focus from concepts, which allegedly underlie uses of verbs of motion, to resources that are provided by verbs of motion for communicating various types of experience. Thus, the proposed description is aimed at covering all frequent types of uses of motion verbs from Table 1. A few exceptions concern idiomatic expressions, like *leaving aside* or *where it left off*, but they are better viewed as lexical items of their own.

### 4.1 Basic options for using verbs of motion

The systemic network that covers verbs referring to away-from motion in English, German and Russian is given in Figure 3. It includes the following options:[5]

**The Type of Motion:** physical vs. abstract

(19) *He left the room* vs. *He left the job*.

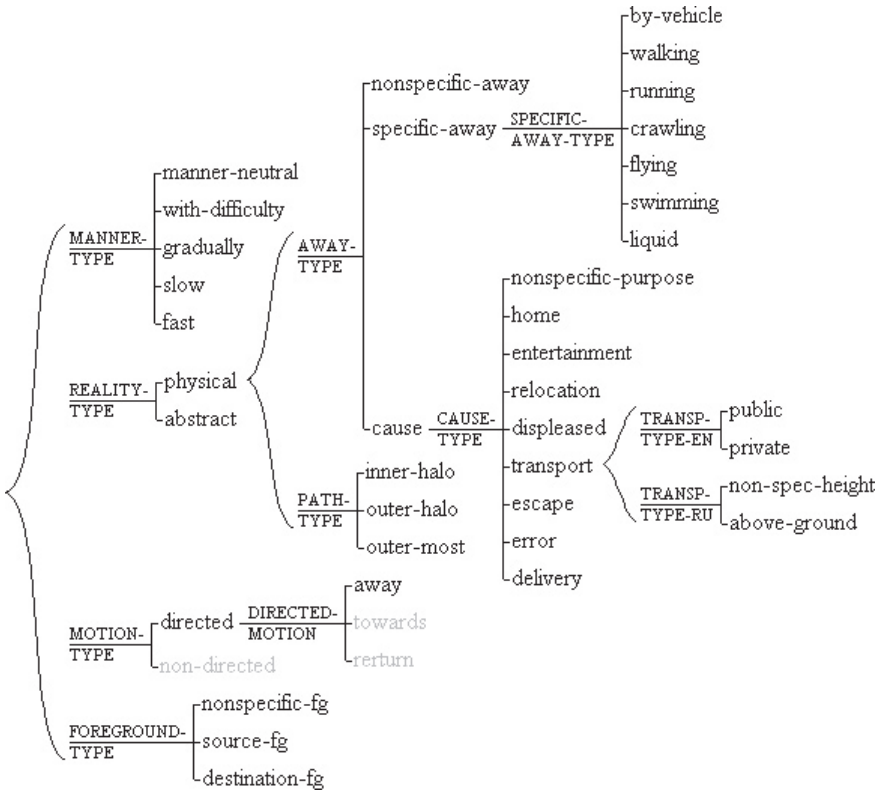


**Figure 3.** The basic options for verbs of motion

The option of abstract motion is considered in Section 4.2 below. In the case of physical motion, there are three possibilities of lexicalisation: [nonspecific-away], a generic reference to motion (*they went to Houston*), [specific-away], a specific reference to the carrier or conduct of motion (*they flew out to Houston*) and [cause], an explicit reference to the cause or purpose of motion (*they moved out to Houston*). Frequent subtypes in the second case are [by-vehicle] (*drive away, wegfahren, uehat'*[6]), [walking] (*go away, weggehen, idti*, when they realise the [specific-away] option), [running] (*run away, weglaufen, ubezhat'*), [crawling] (*creep away, wegkriechen, vypolzti*), etc.

There are important multilingual differences between uses of [nonspecific-away] and [by-vehicle] in English, German and Russian. In English, [nonspecific-away] references like *leave* or *go* (*away*) are considerably more frequent than [by-vehicle] ones, like *drive away*, especially in cases, when the traveler does not drive, but rides in a car, bus or train. In German and Russian, such situations are expressed using forms of *fahren, ehat'* (to go by a vehicle), which

are obligatory for referring to long-distance motion, which necessarily uses a vehicle:

(20) *She has flatly refused to go. [to a summer camp]*
*Sie hat es rundweg abgelehnt wegzufahren.*
*Ona reshiteljno otkazalas' ehat'.*

(21) *I am leaving tonight.*
*Ich fahre heute abend weg.*
*Ja uedu segodnja zhe vecherom.*

As a result, the joint frequency of *drive* and *ride* in English is about 216 instances per million words, while the frequency of the Russian *ehat'* (with respective prefixes) reaches 955 instances per million words.[7]

Non-specific references to motion using a vehicle in German and Russian are possible in very few cases: in particular, when a vehicle departs according to a timetable or follows a determined route: *Avtobus ushel v 9.00* (the bus left at 9.00 according to a timetable) vs. *Avtobus uehal v 9.00* (an occasional bus left at 9.00), see Raxilina (2000). Finally, motion using a vehicle in German and Russian can be expressed in a generic way for habitual destinations, like home, work, school: *Jetzt gehe ich nach Hause, Sejchas ja uhozhu domoj* (Now I am going home).

In addition to the motion type, the three languages also have several lexicalisation options which are specific with respect to the cause of "away-from" motion:

1. [entertainment], to leave home for entertainment, e.g. *go out, ausgehen, vyjti proguljat'sja*;
2. [relocation], to go to live in a different place, e.g. *move away/out, gehen nach X, vyehat'*;
3. [home], to run away from home, often "because you are unhappy there" (CCED), *to run away, von zu Hause weglaufen*, *sbezhat'/ujti iz doma*;
4. [displeased], some motion verbs used in imperative can refer to expressing the anger of the speaker: *get away, get out of my house; geh weg, komme weg; uhodite, ubirajtes'*;
5. [end-journey], to leave a vehicle at the end of one's journey — the language-specific differences have been discussed above: *get out/off, aussteigen, sojti/vyjti*;
6. [escape], to get away from a place of confinement: *escape, entkommen, sbezhat';*

7. [error], to go in a wrong direction; *sich verlaufen, vom Weg abkommen,* also *sich verfahren, verfliegen*, this is an example of a communicative intention, which is frequently realised by verbs of "away-from" motion in one language (German) and by other means in other languages, for instance, by emphasising the result *to get lost, zabludit'sja* in English and Russian. There are two more English verbs: *to wander off* and *to stray* (though they are too infrequent to appear in Table 1), which communicate a different situation: they highlight that the speaker does not know where the 'wanderers' are, whereas *to get lost* highlights that the wanderers themselves do not know where they are (Geoff Thompson, pc). Thus, the most typical subject of *to get lost* is one of the first person pronouns, while the typical subject of *to wander off* and especially *to stray* is a non-interactant.
8. [delivery], to be sent (for a package), most typically as a participle in German: *the letter went, der abgehende Brief, pis'mo ushlo*.

**The Direction of motion:** [non-directed] vs. [towards] vs. [away] vs. [return]

(22) *She ran for two hours vs. She entered the room vs. She left the room.*

A lexical item may either designate an inherently directed motion (*arrive, depart, descend, return*) or be indifferent to its direction (*roll, slide, walk*). In the case of inherently directed motion, it is possible to express the source and the destination, one of which is treated as a reference point, so that the motion is either away from it or towards it (*to enter a room* vs. *to leave a room*). The reference point is the selected point for the Path of motion in Talmy's (1985) terms. Languages also have lexical resources for additional specification of properties of the Path. In the case of physical "away-from" motion, Russian has the following three options for expressing properties of the source and destination:[8]

1. [inner-halo]: X has left a region; this is the unmarked option which is often realised by verbs with the prefix *vy-*, like *vyjti* (go away), *vyehat'* (drive away), *vybezhat'* (run away), *vyletet'* (fly away), etc;
2. [outer-halo]: X has left a region, but remains in the proximity to it; the speaker emphasises the fact that X can return back easily, the option is often realised by verbs with the prefix *ot-* (*otojti, ot'ehat', otbezhat', otletet'*);
3. [outer-most]: X has left a region for a remote destination; the speaker emphasises that X is not expected to return back soon, the option is often realised by verbs with the prefix *u-* (*ujti, uehat', ubezhat', uletet'*).

At the same time, what is considered as [outer-halo] or [outer-most] depends on the current semantic configuration and not solely on properties of the destination. Example (23) from the translation of 'Alice' by Demurova:

(23) *My … uhodili na ulitsu i tselyj den' igrali v klassiki*
'we … went to street and whole day played classics' (a game)[9]

shows that the temporal specification makes it possible to consider the destination (the street) as remote, even though it is clear from context that the place is close to the school. If *vyhodit'* or *othodit'* were used in this situation, this should mean "for a short time", so this would contradict the whole day activity. Even though English and German lack regular lexical means to express [outer-halo] and [outer-most], there are several types of uses that are related to the last options. For instance, CCED lists a sense of *quit* "If you *quit* a place, you leave it completely and do not go back to it".

**The Foreground type:** source- vs. destination-focused

(24) *Forty minutes later they left Colombian airspace* vs.
(25) *This year he has left for St. Louis early to acclimatize himself.*

These examples can illustrate the relationship between intentions, lexicogrammatical options and specific words and constructions. Intentions, such as the emphasis on the destination, are realised by lexicogrammatical options, such as [destination-fg], which are in their turn realised by compatible words and patterns, such as *to leave for X*. In this view, the task of a functional lexicon is to enumerate the most typical ways for expressing [destination-fg].

Unlike the direction of motion, which is more or less fixed for a lexical item (*to leave* cannot realise any physical motion apart from [away-from]), the focal point of the message can be related to specific patterns of uses of the same lexical item:

(26) *They wanted to get out of the country* [escape source-fg] vs.
(27) *If you get out, you go to places and meet people, usually in order to have a more enjoyable life* [entertainment destination-fg].

### 4.2 The subcase of abstract motion

Physical motion provides a schema for conceptualising various events in terms of motion from one point to another. Analyses of "abstract" motion in Langacker (1991) or "fictive" motion in Talmy (1999) show some cases when verbs of motion are used for coding events that do not refer to motion in physical

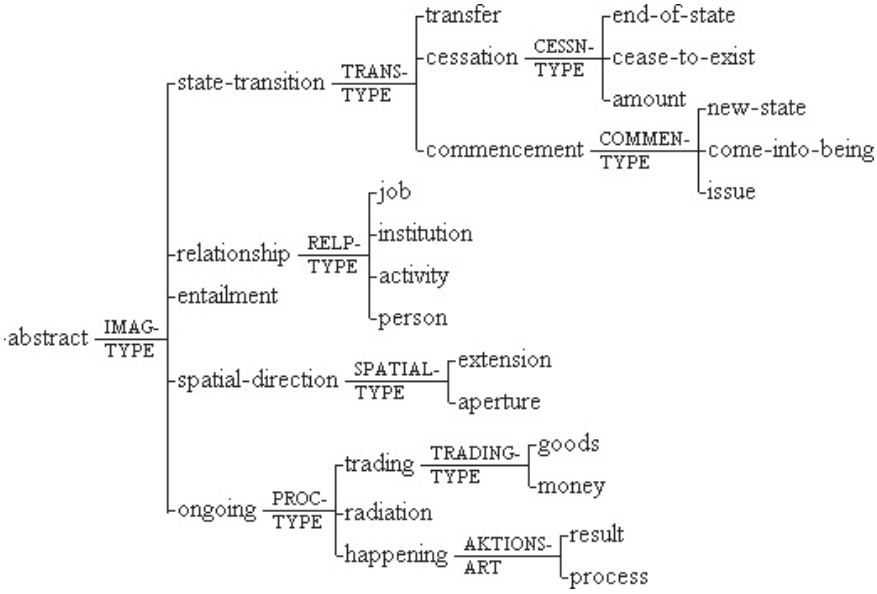


**Figure 4.** Basic options for abstract "away-from" motion in English, German and Russian

space, like *A fence runs along the road.* However, the complete range of such uses is much more diverse: 52 senses in CCED for verbs from Table 1 refer to physical motion, and 63 senses refer to abstract motion; the same situation holds in German (72 vs. 132)[10] and Russian (35 vs. 77). The set of uses ranges from changing social relationships to spreading goods to stopping an activity to designating small amount of resources left. Several German and Russian verbs etymologically related to motion, e.g. *auskommen* or *izbezhat'*, are used in modern language exclusively for referring to abstract motion. However, all uses of abstract motion allow a relatively compact description.

The systemic network shown in Figure 4 aims to cover the most frequent options referring to abstract "away-from" motion in English, German and Russian. Five basic types of abstract "away-from" motion can be recognised: [state-transition], [relationship], [ongoing], [entailment] and [spatial-direction], each of which uses a specific interpretation of the physical model of motion from a source to a destination.

**[state-transition]**

This is a very common type of use, which presents a change of X's state as motion from an initial state to a target state. There are three possibilities:

1. [transfer] — X goes from one state to another: in this case both the source and destination can be in focus. The available lexical choice depends on the nature of the state, e.g. a deliberate topic change: *to leave the topic, vom Thema abkommen, ujti ot temy*; avoiding punishment: *to get off the charge, der Strafe entgehen, izbezhat' nakazanija,* transferring activity: *the job went to X, rabota ushla k X*.
2. [cessation] — X ceases to be in a state: only the source can be in focus. Three options are possible: [end-of-state] — a state ends, this also includes the case when X ceases to function properly, e.g. *a light goes out, die Lampe geht aus, lampa vyshla iz stroja*; [cease-to-exist] — an object or phenomenon disappears, e.g. *a stain came out, der Fleck ist herausgegangen, pjatno soshlo*; [amount] — a resource is used up, e.g. *supplies ran out, vse zapasy ushli*, especially including time and time-related amounts in English and German: *time*/*deadline*/*visa ran out*, *die Zeit*/*Frist*/*Visum ist abgelaufen* (the option in Russian is used in idiomatic expressions: *vremja ushlo, sroki ushli*);
3. [commencement] — X enters a new state: either the source or the destination can be in focus, depending on the perspective. The new state can be presented as the destination of "away from" motion, its causes can be presented as the source. Again, three options are possible here: [new-state] — *to come/go out on strike, in Streik treten, vyjti na rynok/uroven'* (lit. "go out to market/level"), [come-into-being] — an object or phenomenon starts to exist or shows itself, *the sun came out, die Sonne ist herausgekommen, solntse vyshlo*; *the truth came out, die Wahrheit ist herausgekommen, Pravda vyshla na poverhnost'*; [issue] — an information source appears, *the ban/book came out, das Buch ist herausgekommen, zapret/kniga vyshli.*

**[relationship]**

Another very frequent type of using abstract "away from" motion is to designate that a person (X) ends his/her relationships with Y. The following cases are possible:

1. [job] — *to quit one's job, to leave one's office, aus dem Amt gehen, ujti s raboty;*
2. [institution] — *to leave school*, *den Ausschuss verlassen, von der Schule abgehen, vyjti iz komissii, ujti iz shkoly;*
3. [activity] — *to pull out of the project, to withdraw from the talks*, *sich vom Projekt zurückziehen, ujti iz proekta*;

4. [person] — *to leave a person, sbezhat' iz semji* (lit. "run out of one's family"), or *his spouse walked out on him, ihm ist seine Frau weggelaufen, u nego ushla zhena*.

**[entailment]**
A logical relationship between two states of affairs can be present as abstract "away from" motion from one state of affairs to another: *follow* (*from*), *folgen* (*daraus*), *vytekat'*. The choice of the option leads to grammatical consequences, namely, like mental processes, it allows for projection:

(28) *From that it followed logically that he would not be meeting Hildegarde.*
(29) *Vot i vyhodit, chto mne nado ischeznut'*
'so it turns out ("goes out") that I have to disappear'

German and Russian often use this option to express one's assumptions:

(30) *Ich gehe bei meinem Plan davon aus, dass X*
*V svojom plane ja ishozhu iz togo, chto X*
'In my plan I assume ("go out from it") that X.'

**[spatial-direction]**
In this case, motion refers to a particular position of an object;

1. [extension] — a static configuration is expressed using a verb of motion, as discussed by Talmy and Langacker, *a fence runs along the road, der Zaun geht die Bahn entlang, vdol' dorogi idet zabor.* German and Russian verbs of motion are frequently used to express the direction of extension: *der Weg geht nach rechts ab, Doroga uhodit napravo,* 'the road turns (lit. "goes out") to the right';
2. [aperture] — an aperture allows a view/motion towards a certain direction, *Das Fenster geht auf den Garten, Okno vyhodit v sad* (lit. "the window goes to the garden", 'the window faces the garden'); no source can be expressed, the destination is obligatory. In English, the option is quite rare: *The right-hand door goes to a ladder that leads upstairs.*

**[ongoing]**
Yet another possibility to use verbs of motion is to refer to an ongoing activity:

1. [trading] — the feature is actively used in German and Russian and concerns uses of specific motion verbs for evaluation of the dissemination of goods, e.g. *Die Waren gehen gut/schlecht ab*, *Tovary horosho/ploho rashodjatsja* (lit. "the goods go out well/badly", 'they sell well/badly');

2. [radiation] — a psychic quality that can influence other people is presented as radiating out from a person, e.g. *Influence comes from X, Einfluss geht von X aus, Vlijanie ishodit ot X*.
3. [happening] — an event happens (this is most frequently used for evaluation of the process or the resulting state of affairs), e.g. *the interview went well, das Interview ist gut abgelaufen, intervju proshlo uspeshno*. No source or destination is possible.

The last option is used in all three languages, but it is especially frequent in German, which has an array of verbs referring to an event in terms of abstract motion: *abgehen, ablaufen, ausgehen, auskommen, gehen, verlaufen, wegkommen*. The differences between their uses depend either on the subjects of an event (who/what "moves") or on the object of emphasis; in particular, the result or the progress of an event can be emphasised.

The result orientation is realised by *ausgehen* and *wegkommen. Ausgehen* emphasises the result of an event, which can end well or badly, like the result of a game or elections:

(31) *Wenn aber für Helmut Kohl die Wahl doch schlecht ausgehen sollte, …*
'Should, however, the elections for Helmut Kohl end (go out) bad, …'

This often collocates with *am Ende, zum Schluss* (at the end) or *noch mal* (once more).

*Wegkommen* provides an evaluation of the results of an event from the subjective viewpoint, often when they contradict expectations:

(32) *Der Postbeamte ist mit dieser Freiheitsstrafe vorerst gut weggekommen.*
'The post office employee came off with this (judicial) sentence relatively well.'

It can also be used to evaluate one's appearance in a description, like news reports or criticisms, often with the intention of distancing the description from the speaker's opinion:

In contrast to such uses, *ablaufen* and *verlaufen* refer to the progress of a process:[11]

(34) *Alle Spiele verliefen fair.*
"All games ran fair."

(33) *… dass die Deutschen bei Dostojewski ganz wie bei Balzac oder Chesterton schlecht wegkommen,*
'that Germans are shown ("come out") as bad by Dostoyevsky and similarly by Balzac or Chesterton.'

(35) *Wir hoffen, dass das Ostergeschäft ebenso gut abläuft wie der Rest der Saison.*
"we hope that the Easter sale will run out as well as the rest of the season."

They often collocate with *alles* (all), *das Ganze* (whole), or references to time spans *der Rest* (rest), *das Jahr* (year), etc.

*Ablaufen* can be also applied to presentation media (a music record, a film, etc):

(36) *Er versinkt in den Bildern, die wie ein Film in seinem Kopf ablaufen.*
'He immerses himself in the pictures, which run in his head like a film.'

The options for using *gehen* in the sense [happening] are between *ausgehen* and *verlaufen.* Like *ausgehen, gehen* refers to the result of a process and often collocates with *noch mal* (once more), but can also refer to the progress:

*Auskommen* lies also between process and result interpretations. It can be used to emphasise what can be reached with (*mit*) or without (*ohne*) specific means:

(37) *Weil es uns jahrelang zu gut gegangen ist …*
"Since it us year-long too well went, …"
'Since the last year went too well for us, …'
(38) *Da der Motor gut mit sechs Litern auskommt …*
'Since the engine consumes ("comes out well with") six litres'
(39) *Man kann mit 5000 DM für eine funktionstüchtige Küche gut auskommen.*
'One can find ("come out with") a functional kitchen for 5000 DM.'

It often collocates with specifications of amounts of money, effort or functional abilities required for reaching the result. Similarly, *mit X auskommen* refers to the quality of lasting relationships with X:

(40) *Wir sind gut miteinander ausgekommen.*
'We got on ("came out") together well.'

## 4.3 The difference between logic- and communication-centered descriptions

The network shown in Figure 4 is able to accommodate some points of the analysis of the Russian verb *vyjti* from Apresjan (2000), in so far as it is compatible with a description developed in the logic-centered MTT paradigm. The natural question is: what is the difference between meaning-as-concept and meaning-as-use analyses? After all, the ambiguity between physical vs. abstract

motion in *to leave the battle field* (*vyjti iz boja* in Apresjan's original analysis) is still represented as the ambiguity between features in the network. Terminal nodes in the network can be considered as identical to senses, while intermediate nodes can be considered as groups of senses.

The view "from below" maps individual lexical items to possible combinations of features in the network. The structure of the description is similar to the distinction between senses available in dictionaries: a set of features often corresponds to a sense. However, definitions in dictionaries are typically based on concepts, while features in the network correspond to functions for using lexical items. CCED is an example of a dictionary in which definitions are oriented towards possible uses of words rather than definitions of concepts, even though its lexical entries are not formal enough to be suited for computational applications and the requirements imposed by the ECD model. CCED definitions are aimed at human readers, most specifically learners of English. It tries to teach them how to use English lexical items correctly by suggesting situations in which a learner can use a word, for example:

(41) If you *leave* someone to do something, you go away from them so that they do it on their own

The ECD uses another concept of correctness: it aims at the theoretically correct description of lexical items independently from purposes they are used for. As the result of different intentions, definitions differ significantly. Definitions in CCED are informal and circular. For instance, its entries for *go away* and *leave* refer to one another, as is evident when you compare (41) to another definition from CCED:

(42) If you *go away*, you leave a place and spend a period of time somewhere else, especially as a holiday.

This circularity is not allowed in the formal model of ECD, but it may be of help for human readers of CCED, who, most probably, know the basic meaning of *go away* and *leave* and need to know constraints on their uses in specific situations. However, CCED lacks the formal means which would allow the descriptions to be used in computational applications.

The major positive impact of the proposed description is that it makes it possible to account for diverse uses of verbs of motion in a relatively compact way. For 70 verbs of motion in the three languages, the model covers 325 senses taken from respective monolingual dictionaries. This can be achieved because of the shift in perspective: the proposed network does not reify types of uses

corresponding to features as concepts. A "sense" is instantiated when the network is traversed in the context of the current communicative intentions of the speaker (if the network describes the process of natural language production) or in the context of the current attitude of the hearer (if the network describes the process of understanding).

This allows for greater granularity and flexibility in adjusting uses to their context. With respect to granularity, the proposed network describes different resources used for leaving a vehicle at the end of the journey, while WordNet provides a concept (synset) "move out of; as of a room, a country, a bus, etc" which joins an open list of apparently different concepts of leaving. Verbs belonging to the synset (*exit, go out, get out, leave*) cannot be used in all situations as synonyms: ?*He exited the country,* **He went out from the bus,* ?*He got out from the room.* On the other hand, it is clear that the designers of WordNet tried to reduce proliferation of such concepts as LEAVE-COUNTRY, LEAVE-BUS, LEAVE-CAR. This problem does not appear when lexical items are considered as resources, so that respective features (like [end-journey], [public/private] transportation means, etc.) are introduced when it is necessary to reflect the difference in the lexical choice.

With respect to flexibility, the possible ambiguity is dynamic. For instance, in the case of *vyjti iz boja* (*to leave the battle field*), one can consider two possible traversals or prefer one of them depending on the context without the requirement for two separate concepts representing two senses. If a verb of motion is used in a particular context, the network provides the possibility of linking its use via lexical features to communicative intentions underlying the use. Even though the Generative Lexicon model also aims at a greater flexibility of its lexical descriptions, it is unlikely that the fixed inventory of semantic roles provided in the model can achieve the same result in accounting for all most frequent ways of using verbs of motion as described above. At least, all available descriptions of lexical items, as in Pustejovsky (1995), cover only a small subset of word uses.

### 4.4 The preservation of lexical features in translations

The next natural question is: how multilingual is the network presented? The majority of general options for expressing events of different types as motion are strikingly similar in the three languages. A few options that are significantly different are discussed above explicitly. However, the question is whether or not identical situations are expressed by the same set of resources.

**Table 2.** The distribution of motion types in translations of "Alice in Wonderland"

| Corpora | Physical motion | | | | | Abstract motion | | | | | Exclamation | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | P=P | P=M | P=X | P=S | P=0 | A=A | A=X | A=M | A=S | A=0 | E=E | E=M | E=0 |
| Carroll | 182 | | | | | 130 | | | | | 9 | | |
| Demurova | 132 | 21 | 7 | 6 | 16 | 23 | 8 | 85 | 2 | 12 | 8 | 0 | 1 |
| Nabokov | 130 | 21 | 12 | 4 | 15 | 28 | 13 | 71 | 3 | 15 | 8 | 0 | 1 |
| Zahoder | 119 | 23 | 13 | 0 | 27 | 24 | 13 | 77 | 3 | 13 | 8 | 1 | 0 |
| Average Ru | 127 | 21.7 | 10.7 | 3.3 | 19.3 | 25.0 | 11.3 | 77.7 | 2.7 | 13.3 | 8 | 0.3 | 0.7 |
| Zimmer-mann | 153 | 11 | 7 | 1 | 10 | 62 | 9 | 55 | 0 | 4 | 8 | 1 | 0 |

An aligned parallel corpus provides the possibility of checking which features of verbs of motion are kept in translations. The study is based on uses of verbs of motion in *Alice in Wonderland*, three translations into Russian and one into German. The corpus is small and specialised, however, its size allows all uses of verbs of motion in the three languages to be checked. The English original has in total 333 uses of verbs of motion, 182 of which express physical motion, while 130 refer to abstract motion.[12]

The distribution of possible cases for translations is depicted in Table 2. Table 2 should be read as follows:

- the translation of motion is also expressed by motion of the respective type, these are columns P=P or A=A in the table ([physical] is expressed by [physical], [abstract] by [abstract]);
- the translation refers to a process of another type, the cases here are [P] or [A] = M (material), X (existential) or S (mental, i.e. saying or sensing);
- a reference to the event is omitted from the translation, P=0 or A=0;
- a verb of motion used for an exclamation, e.g. *Come, there's half my plan done*. This can be translated as an exclamation (E=E), differently (E=M) or omitted (E=0);

The last option is quite infrequent and its function is different from those analysed above, so we can focus on expressions of [physical] and [abstract] motion. In the corpus, physical motion in English is typically expressed by physical motion (the P=P column) in both German (153 cases, 84%) and Russian (127 cases on average, 70%). Less often verbs of motion are rendered by processes of another type, for example, by other material processes (P=M: 12% in Russian, 6% in German). Compare (43) to its German translation:

(43) *tell me, please, which way I ought to go from here?*

(44) *willst du mir wohl sagen, welchen Weg ich hier nehmen muss?*
'would you tell me which way should I take'

English motion verbs can also be reported in a more generic way: 6% of references to physical motion are translated as existential processes in Russian, 4% in German. For instance:

(45) *After these came the royal children*
(46) *Presently the Rabbit came up to the door*

are translated respectively as:

(47) *Zatem pojavilis' korolevskie deti*
then appeared royal children
(48) *Krolik byl uzhe u dverej*
Rabbit was already at doors

These expressions refer not to motion, but to its result. In this case, the properties of the process, e.g. its destination, are expressed within the circumstantial subtype of relational processes. Relatively often (11% in Russian, 5% in German) a reference to physical motion is simply omitted from the translation.

In short, the feature [physical] is often preserved in translation. Some more delicate features of physical motion are also preserved, when the target language has lexical items with necessary properties and these lexical items can be realised in context. However, a translator may emphasise a specific property of motion. For instance, the sentence (49) and its respective translations are described by the following sets of features:

(49) *was just going to leave the room.*
[directed away source-fg physical nonspecific-away]
(50) *sovsem uzhe sobralas' vyjti iz komnatki*
just prepared leave-INF from room
[directed away source-fg physical nonspecific-away inner-halo]
(51) *Ona uzhe sobralas' bezhat' obratno,*
she already prepared run-INF back
[directed away destination-fg physical specific-away run-away]
(52) *i sobralas' uzhe idti,*
and prepared already go-INF
[non-directed physical nonspecific-manner]
(53) *und wollte eben das Zimmer verlassen,*
and wanted just the room leave-INF
[directed away source-fg physical nonspecific-away]

One translation (52) uses [non-directed] motion and does not specify the Path. Another translation (51) foregrounds the destination and expresses the manner of motion explicitly (as running). Other translations (50), (53) keep the set of features used in the sources. Nevertheless, the general setup of motion is preserved.

Types of references to abstract motion are not only diverse, but also quite frequent. As Table 2 shows, there are 151 cases where a verb of motion in 'Alice' does not refer to a motion event; this is comparable to 182 references to physical motion. In other registers, like newspapers or science texts, the number of abstract motion uses is even higher than uses of physical motion. For instance, in a random sample of 100 concordance lines of the newspaper section of the British National Corpus queried for English verbs from Table 1, only 23 instances use motion verbs in the physical sense. With respect to translations, the set of features is less stable: 19% of abstract motion in English is translated into Russian as motion (the figure for German is larger: 48%). Most typically (60% in Russian vs. 42% in German), another type of process is used:

(54) *Alice soon came to the conclusion that …*
(55) *Alisa skoro reshila, chto …*
Alice soon decided that

However, the opposite is also true: abstract motion in German and Russian is frequently used in situations where English uses processes of another type. For instance, example (4) is a translation of *it would have made a dreadfully ugly child.*

## 5. Conclusions

The aim of the study reported in the paper is to contrast models for describing lexical meanings in a formal framework which assumes that words are pointers to senses, and in a functional framework which assumes that words are resources and are used to realise communicative intentions of the speaker. The corpus-based study of uses of verbs of motion illustrates that even though lexicographic research aimed at enumerating senses of lexical items provides an invaluable resource for describing typical meanings, the theoretical model of word uses cannot be based on the notion of a dictionary-like mapping between words and senses, because real uses of verbs in texts often violate the necessary and sufficient conditions for treating them as instances of specific

concepts: uses often refer to slightly different concepts or refer to several concepts simultaneously.

At the same time, the communication-oriented model can achieve a more concise and robust description of basic types of uses for sets of lexical items in several languages, by using the systemic network for relating communicative intentions to lexical items. The objective of the case study of verbs of motion follows Halliday's aim to construct a grammar "that would make it possible to say sensible and useful things about any text, spoken or written, in modern English" (Halliday, 1985:xv). Similarly, my objective in the study was to say sensible and useful things about the most frequent uses of the most frequent verbs of "away-from" motion in English, German and Russian, i.e. to show what resources they provide for communicating various types of events in the three languages. This resulted in a lexicographical database covering 430 types of uses of 70 headwords for the three languages by means of 19 systems with 66 features.[13]

Leaf nodes in the network of lexical features may look similar to senses in WordNet or ECD. However, the crucial difference of the lexical systemic network compared to the dictionary-based approach is that sets of features do not constitute concepts. The network specifies the potential for using words. The potential is instantiated by a traversal of the network in context, but the traversal does not lead to reification of a concept to which a word points. This means that the meaning of a lexical item in the context depends on its contribution to the ongoing discourse. The systemic network can also address the issue of similarities and differences in uses of verbs of motion in English, German and Russian without the need to map between language-specific concepts.

## Notes

* The study reported in the paper was supported by the Alexander von Humboldt Fellowship for 2001–2002. I am most grateful to various persons with whom I discussed the ideas presented here, primarily to Tony Hartley, who devoted a lot of effort to improving the paper, as well as to John Bateman, Irina Kobozeva, Lena Sokolova and Constanze Vorwerg. The coding and the statistical counts have been done using Mick O'Donnell's Systemic Coder, http://www.wagsoft.com.

**1.** It is considered in the example by Lyne (1988): Do you think that the bike is pushable?

**2.** The notion of intention used here refers to any piece of semantic information that is going to be expressed by the speaker, thus including not only conscious intentions, but also such distinctions as those between types of carriers in riding, flying, swimming, etc.

**3.** A very few exceptions concern idiomatic expressions, like *Ein Zug ging um 03.00 Uhr in Richtung Westen ab.* ('A train went at 3pm to the West', with respect to a special train for refugees).

**4.** For another implementation of the lexical resources as a separate stratum, cf. Stede (1999).

**5.** There are other options that influence lexical meanings of verbs of motion, e.g. activity vs. achievement (Levin, 1993). However, they are not specific for verbs of motion, so their choice should be represented in less delicate portions of the network.

**6.** The order of all examples is English, German, Russian, unless explicitly specified. In almost all cases the examples are equivalent in meaning.

**7.** The frequency data are from the BNC for English and the Russian Reference Corpus for Russian.

**8.** The names for the three options follow the analysis of French verbs of motion by Asher and Sablayrolles (1995).

**9.** Note that the original is different at this point, because this is a translation of a pun: *I went to the Classics master … he taught Laughing and Grief.*

**10.** The number of senses in German is large partly because of the wide range of uses of German motion verbs and partly because of a very elaborate system of senses used in the German Wahrig dictionary.

**11.** *Abgehen* is often used for this purpose in Southern German dialects.

**12.** The study is not restricted to verbs of "away from" motion from Table 1. The count in Table 2 includes all frequent verbs of motion listed in Miller (1972) and their translations.

**13.** The design of the database is discussed in Sharoff (2002).

## References

Apresjan, J.D. (1974) Regular polysemy. *Linguistics* 142: 5–32.

Apresjan, J.D. (2000) *Systematic lexicography*. Oxford: Oxford University Press.

Asher, N., and Sablayrolles, P. (1995). A Typology and Discourse Semantics for Motion Verbs and Spatial PPs in French. *Journal of Semantics* 12: 163–209.

Bateman, J.A., Teich, E., Kruijff G-J, Kruijff-Korbayova I., Sharoff S. and Skoumalova H. (2000) Resources for multilingual text generation in three Slavic languages. In *Proc. Languages Resources and Evaluation Conference LREC2000,* Athens, Greece, May 30-June 2, 2000, pp. 1763–1767.

Bierwisch, M. and Schreuder, R. (1992) From concepts to lexical items. *Cognition* 42:23–60.

CCED, 1995. *Collins Cobuild English dictionary.* 2nd edition. Glasgow: HarperCollins.

Eichinger, L.M. (1989). *Raum und Zeit im Verbwortschatz des Deutschen: eine valenzgrammatische Studie*. Tübingen: Niemeyer.

Fillmore, C.J. (1983). How to know whether you're coming or going. In G. Rauch (ed.) *Essays on deixis*. Tübingen: Narr. 219–227.

Gerling, M. and Orthen, N. (1979) *Deutsche Zustands- und Bewegungsverben: eine Untersuchung zu ihrer semantischen Struktur und Valenz.* Tübingen: Narr.

Grenoble, L.A. (1986) A contrastive analysis of verbs of motion in Russian and Polish. PhD Thesis, University of California, Berkeley.

Halliday, M.A.K. (1961) Categories of the theory of grammar. *Word,* 17:241–292.

Halliday, M.A.K. (1985) *Introduction to functional grammar.* London: Edward Arnold.

Halliday, M.A.K. and Matthiessen, C.M.I.M. (1999) *Construing experience through meaning: A language-based approach to cognition.* London: Cassell.

Hasan, R. (1987) The grammarian's dream: lexis as most delicate grammar. In M.A.K. Halliday and R.P. Fawcett (eds.) *New developments in systemic linguistics.* Volume 1. London: Pinter Publishers. 184–211.

Hunston, S. and Francis, G. (1999) *Pattern grammar: A corpus driven approach to the lexical grammar of English*. Amsterdam: John Benjamins.

Kilgarriff, A. (1997) "I don't believe in word senses". *Computers and the Humanities* 31 (2): 91–113. Also at http://ftp.itri.bton.ac.uk/pub/reports/ITRI–97–12.ps.gz

Langacker, R. (1988) A usage-based model. In B. Rudzka-Ostyn (ed.) *Topics in cognitive linguistics.* Amsterdam: John Benjamins. 127–161.

Langacker, R. (1991) *Concept, image, and symbol: The cognitive basis of grammar.* Berlin: Mouton de Gruyter.

Levin, B. (1993). *English Verb Classes and Alternations: A Preliminary Investigation*, Chicago: The University of Chicago Press.

Levin, B. and Rappaport Hovav, M. (1992) The lexical semantics of verbs of motion: The perspective from unaccusativity. In I. M. Roca (ed.) *Thematic structure: Its role in grammar.* Berlin: Walter de Gruyter. 247–269.

Lyne, A.A. (1988) Systemic syntax from a lexical point of view. In J. Benson, M. Cummings and W. Greaves (eds.) *Linguistics in a systemic perspective*. Amsterdam: John Benjamins. 53–72.

Mackenzie, I. and Mel'chuk I. (1988) Crossroads of obstetrics and lexicography: a case study. *International Journal of Lexicography* 1: 71–83.

Martin. J.R. (1987) The meaning of features in systemic linguistics. In M.A.K. Halliday and R.P. Fawcett (eds.) *New developments in systemic linguistics.* Volume 1. London: Pinter Publishers. 14–40.

Matthiessen, C.M.I.M. and Bateman, J.A. (1991) *Text generation and systemic functional linguistics: Experiences from English and Japanese.* London: Pinter Publishers.

Mel'chuk, I.A. (1988) Semantic description of lexical units in an Explanatory Combinatorial Dictionary: basic principles and heuristic criteria. *International Journal of Lexicography* 1: 165–188.

Miller, G. (1972) English verbs of motion: a case study in semantics and lexical memory. In A.W. Melton and E. Martin (eds.) *Coding processes in human memory.* New York: Wiley. 335–372.

Miller, G. (ed.) (1990) WordNet: an online lexical database. *International Journal of Lexicography* 3 (special issue on WordNet). The issue is also available as: *Five papers on WordNet* at http://www.cogsci.princeton.edu/~wn/obtain/5papers.pdf

Pustejovsky, J. (1995) *The generative lexicon.* Cambridge, MA: MIT Press.

Raxilina, E, (2000) *Kognitivnyj analiz predmetnyh imen: Semantika i sochetaemost'*. Moscow: Russkie Slovari.

Sharoff, S. (2002). Meaning as use: exploitation of aligned corpora for the contrastive study of lexical semantics. In *Proc. of Language Resources and Evaluation Conference (LREC02).* May, 2002, Las Palmas, Spain. 447–452.

Sowa, J.F. (1991). Towards the expressive power of natural language. In J.F. Sowa (ed.) *Principles of semantic networks.* San Mateo, CA: Kaufman, 1991. 157–189.

Stede, M. (1999). *Lexical semantics and knowledge representation in multilingual sentence generation*. Boston: Kluwer.

Talmy, L. (1972) Semantic structures in English and Atsugewi. PhD Thesis, University of California, Berkeley.

Talmy, L. (1985) Lexicalization patterns: semantic structure in lexical forms. In T. Shopen (ed.) *Language typology and syntactic description, vol. 3: Grammatical categories and the lexicon*. Cambridge: Cambridge University Press. 57–149.

Talmy, L. (1999) Fictive motion in language and 'ception'. In P. Bloom, M.A. Peterson, L. Nadel and M.F. Garrett (eds.) *Language and space.* Cambridge, MA: The MIT Press.

Tucker, G.H. (1998) *The lexicogrammar of adjectives: a systemic functional approach to lexis.* London: Cassell.

Wanner L. (1997) Exploring lexical resources for text generation in a systemic functional language model. PhD Thesis, University of Saarbrücken.

Winograd, T. (1976) Towards a procedural understanding of semantics. *Revue Internationale de Philosophie* 30: 260–303.

### *Author's address*

Serge Sharoff
Centre for Translation Studies
University of Leeds
Leeds LS2 9JT
UK

s.sharoff@leeds.ac.uk

## Appendix 1. Senses of LEAVE in WordNet 1.6 (http://www.cogsci.princeton.edu/~wn/)

The verb LEAVE has 17 senses in WordNet

1. **leave**, go forth, go away — (go away from a place; "At what time does your train leave?" "She didn't leave until midnight")

2. **leave** — (leave behind; "She left a mess when she moved out"; "His good luck finally left him")
3. **leave** — (cause to be in a specified state; "The inflation left them penniless"; "He left money on the table for his wife to find")
4. **leave** — (leave unchanged or unaltered; "leave it as is")
5. exit, go out, get out, **leave** — (move out of; as of a room, a country, a bus, etc.)
6. **leave**, allow for, allow, provide for — (make a possibility or provide opportunity for; "This leaves no room for improvement"; "The evidence allows only one conclusion")
7. **leave**, result, lead — (result in; "The water left a mark on the silk dress"; "Her blood left a stain on the napkin")
8. **leave**, depart — (leave home, school, a position, etc.)
9. **leave**, let — (let be; leave alone or undisturbed; "leave the door open!")
10. entrust, **leave** — (put into the care or protection of someone; "He left the decision to his deputy")
11. **leave**, leave behind — (refrain from taking; "Please leave the hand-outs on the tables")
12. bequeath, will, **leave** — (leave or give by will; "My aunt bequeathed me all her jewelry")
13. **leave** — (have left or have as a remainder; "That left the four of us"; "19 — 8 leaves 11")
14. **leave**, leave behind — (be survived by after one's death: "He left six children")
15. impart, **leave**, give, pass on — (tell, give knowledge; "give a secret to the Russians")
16. forget, **leave** — (leave behind; "I forgot my umbrella in the restaurant")
17. drop out, give up, throw in, throw in the towel, quit, **leave**, chuck up the sponge — (give up in the face of defeat of lacking hope; admit defeat; "In the second round, the challenger gave up")

## Appendix 2. Senses of VYJTI from (Apresjan, 2000: 261–278)

1.1 'to move out by foot': 'Living being A, on foot, moved out of enclosed space B into a more open space C'; *vyjti iz komnaty* (to leave the room);
1.2 'to move outside': 'Vehicle A moved out of enclosed space B into a more open space C'; *Sudno vyshlo iz buxty* (The ship left the harbor);
1.3 'to go away from a place': 'People or vehicle A, located at point X and intending to reach point Y began moving towards Y and are located on the route from X to Y'; *Polk vyxodit zavtra* (The regiment leaves tomorrow);
2. 'to finish one's service or location somewhere': 'Person A ceased to be situated in institution X, where he had been confined for treatment or punishment'; *vyjti iz tjur'my* (to leave the prison);
3.1 'to cease to be a part or a member': 'Person A, not wishing to remain a member of organization or group X, took official steps to cease being a member and ceased being a member of X'; *vyjti iz sostava komissii* (to leave the committee);
3.2 'to cease to be in a state': 'Object A ceased to be in state X'; *vyjti iz povinovenija* (to get out of control);

3.3 'to cease to do': 'Person or people A ceased to perform action X'; *vyjti iz povinovenija boja* (to leave the field of battle);
4. 'to come to en end, to be used up': 'In the course of some human activity a certain amount of resource X was used up'; *Za mesjats vyshlo okolo kubometra drov* (about as cubic meter of firewood was used up in a month);
5.1 'to come, to appear': 'Person A appeared at his place of usual activity B after a certain interval'; *vyjti na rabotu* (to come to work);
5.2 'to appear, to become accessible': 'Object A was made public and became accessible to consumers'; *vyjti na ekrany* (to be released);
5.3 'to receive access': 'Person A seeking to resolve a matter that can only be resolved by a person holding a high position in the hierarchy of power, gained access to person X'; *vyjti na predsedatelja* (to obtain access to the chairman);
6. 'to enter into matrimony': 'Woman A entered into matrimony with man X'; *vyjti zamuzh* (to get married [of a woman]);
7.1 'to turn into, to become': 'Person X, previously in the class of people X, by virtue of his success, has entered the class of people Y, whose social status is higher than that of class X'; *vyjti v generaly* (to become a general);
7.2 'to have the makings of': 'Person X, whose qualities were such that with training and experience he might become a professional of class A'; *iz nego vyjdet general* (he has the makings of a general);
7.3 'to come into existence': 'From object X, whose properties and quantity were such that an object of class A could be made from it, an object of class A was made'; *iz etoj zatei nichego ne vyjdet* (nothing will come out of this venture);
7.4 'to turn out to be': 'Object A could have several properties different from property X; as a result of somebody's efforts, A acquired the property X'; *vstrecha vyshla interesnoj* (the meeting turned out to be interesting);
8.1 'to result': 'Event of situation X need not have occurred; as a result of somebody's activity X did occur'; *vyshla neprijatnost'* (some unpleasantness resulted);
8.2 'to turn out, transpire': 'A chain of deductions exists, the concluding point of which is a judgement P'; *vyxodit, chto Vy pravy* (it turns out that you are right);
9. 'to face a certain direction': 'Structure A, or part of it, or aperture or exit from it, faces towards object B'; *Okna vyxodjat v sad* (The windows look out over the garden);