Chapter 7

Basic verb vocabulary: An empirical approach to argument structure and word associations

Valentina Stefanova

Department of Computational Linguistics, Institute for Bulgarian Language, Bulgarian Academy of Sciences

Maria A. Todorova

Department of Computational Linguistics, Institute for Bulgarian Language, Bulgarian Academy of Sciences

Tsvetana Dimitrova

Department of Computational Linguistics, Institute for Bulgarian Language, Bulgarian Academy of Sciences

The article discusses the results of a pilot survey on basic verb vocabulary conducted through an online experiment in the form of language tasks employing associative stimuli, thematic stimuli, and contextual and textual stimuli. An analysis revolves around observations on the performance of the tasks by two focus groups of children – 7- to 10-year-olds and 11- to 14-year-olds. Semantic frame representations of the selected verbs are employed to explain the specifics of the respondents' competence in recognising the selected verb meanings and the arguments associated with a specific verb meaning.

1 Introduction

The article discusses the results of a survey on basic verb vocabulary in Bulgarian conducted via an online experiment in the form of five types of language

Valentina Stefanova, Maria A. Todorova & Tsvetana Dimitrova. 2025. Basic verb vocabulary: An empirical approach to argument structure and word associations. In Svetla Koeva (ed.), *Universality of semantic frames and language-specific Bulgarian data*, 299–335. Berlin: Language Science Press. DOI: 10.5281/zenodo.15682196

tasks among two focus groups – children from the primary stage of education (7- to 10-year-olds and 11- to 14-year-olds). The analysis and observations on the performance of the tasks reveal the respondents' basic competences for understanding the selected verb meanings and confirms or rejects hypothesis of basic vocabulary set.

When formulating language tasks and selecting the lexical information they test, we assumed that the respondents of all focus groups had acquired basic vocabulary in their educational and family environment and could understand and use a set of words and expressions associated with different spheres of human activity. The observed data that was part of the experiment was selected following two large language resources – (Princeton) WordNet (Fellbaum 1990) and the Bulgarian wordnet (BulNet) (Koeva 2021a).

The discussion in the article revolves around the verbs with the highest selection frequency by respondents, which we analyse according to their semantic frame representations to shed light on the realisation of specific meanings. The observations are organised around the description of the general structure of the target verb groups, semantic frame elements. Based on these observations, conclusions are drawn as to which semantic frames can describe the verbs.

Our work draws upon the hypothesis that children in the focus age groups have mastered the basic vocabulary, which includes lexical units that are freely and spontaneously used in everyday language practice. Children have been exposed to this part of the basic vocabulary in the course of their schooling and in their family environment.

Firstly, we attempt at confirming the validity of the hypothesis for the target verbs by measuring the frequency of the respondents' answers. Secondly, our aim is to determine the semantic frames which may be evoked by the target verbs, via observations on the specific frame elements recognised by the children to confirm their understanding of the verbs' meaning. The respondent's knowledge of the selected verbs is analysed through Frame Semantics (Fillmore 1982) and the semantic-syntactic description of the concrete frames in FrameNet.

The pilot experiment involved language tasks, described in more detail in Section 3. The tasks within this pilot study aim at examining whether (and which of) the verbs from the set belong to the basic vocabulary, on one hand, and on the other hand – whether the form of the experiment (and specific tasks) is appropriate for testing the respondents' (children of relevant school age) linguistic knowledge, experience and intuition about the use of the verbs, and the ability to match them with the appropriate elements associated with them.

The article is organised as follows: in Section 2, we discuss some definitions along with our motivation; Section 3 gives an outline of the experiment which

has been more thoroughly described in previous papers; Section 4 discusses the structures of the basic vocabulary verbs that have been investigated via the language tasks (the discussion first focuses on the thematic tasks, followed by the association tasks, and the contextual tasks.).

2 Related work and motivation

Basic vocabulary build-up and acquisition has been addressed in numerous linguistic works in recent decades with a solid theoretical as well as applied background (Meara 1980, Arnaud & Bejoint 1992, Schmitt 2007, Cohen 1990).

The process of language acquisition of a person is related to both communicative abilities and skills in mastering different levels and areas of linguistic knowledge like understanding and producing sentences and longer pieces of narrative (Karmiloff & Karmiloff-Smith 2001, Vulchanova & Vulchanov 2021).

The core of the lexical system is organised around active vocabulary (Georgiev & Duridanov 1959, Boyadzhiev 2002) consisting of lexical units that are freely and spontaneously used in everyday language practice, in all spheres of life, throughout a variety of styles, in both oral and written speech (Boyadzhiev 2002). The words that belong to the basic vocabulary are known by all native speakers of a language, serve to derive other words, and are usually inherited from earlier language stages. The basic vocabulary includes nouns denoting basic activities and states, such as eating, drinking, sleeping, lying, sitting, etc. (Georgiev & Duridanov 1959).

There are various studies and experiments of first language acquisition mechanisms in different conditions and at different ages (Vulchanova & Vulchanov 2021, Stoianova 2014, Stenger & Avgustinova 2021). The applied approaches to the study of the basic vocabulary in Bulgarian are only a few most of them involved mainly children in preschool age (Popova 2020, Andonova 2021) and people learning Bulgarian as a foreign language (Dimchev 2005, Nisheva 2013, Burov 2000).

As far as we know, the language competence and vocabulary of Bulgarian children during the first school years, in the initial stage of education, 1st – 4th grade (between the ages of 6 and 11), which are the focus of the interest of this article, haven't been subject of a separate study. According to Vulchanova et al. (2020: 682) this is the second stage of children language learning when words are integrated it into the network of lexical representations. We focus on the period of the first school years, also because it is considered that knowledge about words, grammar and discourse uses are built as independent components

of language competence, which can already be identified separately (Vulchanova & Vulchanov 2021).

In recent decades, strong focus has been placed on technologies for language learning and teaching. The methods used by applied linguistics (Carter 1998) and works on task-based language teaching (Dalpanagioti 2021, Dolgova & Tyler 2019) emphasise the usage-based nature of linguistic knowledge and language and give opportunity for different angles of language investigations.

Furthermore, various studies on language acquisition have employed approaches to language learning and teaching via computationally represented contextual information and the theory of Frame Semantics and FrameNet. Jódar-Sánchez (2018) demonstrate the application of Frame Semantics' principles for learning second language lexis as well as for better understanding and/or acquisition of the learner's first language lexicon. Blanco-Carrión (2006) underlines that frames, being concepts, were organised around human experience.

Employing the theory of Frame Semantics Torrent et al. (2022) consider that the notion of frame includes context as one possible source of information in language comprehension. They report on two experiments: (i) the identification of frame-evoking lexical units in sentences, and (ii) a methodology for domain adaptation in Neural Machine Translation that leverages frames and qualia for representing sentence-level context.

Although technologies are widely used in education in Bulgaria, the Bulgarian basic vocabulary and learner's competence with respect to employing available language resource have not been researched so far.

3 The experiment

The experiment was carried out in the form of language games in an online environment.[1] Each game variant includes a combination of the language tasks, which are described in each section below with a discussion on results. The tasks use different types of stimuli – verbal (target verbs), associative (images), and contextual (grammatical and selective) ones. The verb stimuli are associated with thematic areas related to universal human needs and basic human activities such as nutrition, body movement, daily household activities, personal interests – leisure, favourite pastimes; knowledge of the surrounding world (weather, seasons, climate), plants and animals, etc. (thematic areas are also covered in the WordNet semantic classes (Miller et al. 1990)). Picture stimuli feature clear and recognisable objects that can be associated with actions and states represented

[1]https://ibl.bas.bg/igrasglagoli/

by the verbs. They were selected from databases of free images, and a set of 750 graphical images Duabeitia et al. (2018).

We analyse respondents' knowledge of verb meaning according to the results in the execution of the tasks (association tasks (types 1 and 2), thematic tasks (types 3 and 4), and a contextual task (type 5), discussed in detail in Section 4.

In the two association tasks, the respondents had to choose (at least) one or more verbs according to their association with a picture stimulus and/or simple text stimuli (words, phrases, sentences). In the thematic tasks, the respondents had to choose five out of ten verbs, based only on their membership in a specific semantic class (according to WordNet). In the contextual task, the respondents had to fill in the gaps in a short text by choosing from a list of verbs, given below, and taking into consideration the contextual, grammatical, and selection specifics of the sentence and of the text as a whole.

We compare the respondent's answers with the semantic frames of the target verbs to evaluate the respondents' competence of the situational context and to observe their reactions to linguistic context and accurate use of language units and forms.

The results from the experiment were collected in Excel tables, and the answers were calculated as a percentage of 100%. For example, in a task with four possible answers (verbs) (in task type 1), the results had the following distribution: *hug* 34/82.90%, *caress* 2/4.90%, *cuddle* 4/9.80%, *squeeze* 1/2.40% (number of responses / percentage of all responses given). We determine whether the hypothesis that a verb belongs to the basic vocabulary is correct or not based on the percentage of the answers given. We have used correlation scores to analyse some of the results of the thematic and association tasks.

3.1 The experiment dataset and the target verbs

The verbs used in the experiment were selected from a set of verbs compiled according to basic vocabulary criteria as described in Koeva & Doychev (2022):

(1) the verb's place in the structure of the Bulgarian WordNet (e.g., synonymy and hyponymy-hypernymy relations between verb synonym sets);

(2) the membership of synonym sets to the subsets of base concepts (BCS) in WordNet;[2]

[2]The set of Base concept synsets has been defined by the EuroWordNet and the BalkaNet projects: http://globalwordnet.org/resources/gwa-base-concepts/

(3) the proximity of the synonym set to the root of the lexical semantic tree with respect to the hierarchical substructure in WordNet in which the corresponding synonym set is included;

(4) the frequency of the verbs in the Bulgarian National Corpus (BulNC) (Koeva et al. 2012) (per 1 million words) in texts of different domains and genres;

(5) the membership of the verbs in a list of meanings with an estimate of their age of acquisition according to the study "Test-based age-of-acquisition norms for 44 thousand English word meanings" (Brysbaert & Biemiller 2017).

Additional criteria for the selection of verbs for children are:

(6) verbs belonging to selected semantic classes (Table 1) in WordNet (Miller et al. 1990);

(7) the frequency of the verbs in a small corpus of textbooks for 7-year-olds children on Bulgarian language, fine arts, music, technology.

For example, verbs as *наранявам* 'injure',[3] *отслабвам* 'reduce/lose weight', *режа* 'cut' are selected as target verbs because of their high frequency in BulNC; they belong to the set of BCS in WordNet, and are either at or near the root of the WordNet lexical-semantic tree; they are also found in the 4th grade textbooks.

In addition, we had to take into account the relatively short attention span of children and set some limits on the number and the variety of verbs that were selected in the tasks. Thus, we have selected 188 verbs, of which 112 were targeted in the tasks (considered the "correct" ones).

To provide coverage of the basic vocabulary requirement for inclusion of everyday activities and states, such as eating, drinking, sleeping, lying, sitting, etc., we have manually selected target verbs, using the information of semantic class of verbs in WordNet, distributed as shown in Table 1.

Verbs such as verbs of communication – *write, indicate, teach, pause, hum*; verbs of cognition – *surprise, think, decide, read*; stative verbs – *deserve, miss, be, wait*, are less covered, with few examples of creation verbs (*play*) and emotion verbs (*desire*).

The tasks involve different verb classes (instantiating different frames) which may evoke specific elements associated with them. In associative tasks, verbs of

[3]The Bulgarian examples are followed by their translation equivalents.

Table 1: The distribution of target verbs within WordNet semantic classes

Semantic Class	Number	Type of Tasks[a]	Examples.
verb.contact	34	Associative Tasks	*вися* 'hang', *изрязвам* 'carve', *копая* 'dig', *клякам* 'squat', *намазвам* 'spread', etc.
verb.motion	39	Associative and Contextual Tasks	*бягам* 'run', *излитам* 'take off', *скачам* 'jump', *навеждам се* 'bend', *сядам* 'sit down', etc.
verb.consumption	10	Associative, Thematic Tasks	*дъвча* 'chew', *закусвам* 'have breakfast', *пия* 'drink', *ям* 'eat', etc.
verb.body	11	Thematic and Contextual Tasks	*дишам* 'breathe', *ранявам* 'injure', *поруменявам* 'blush', *спя* 'sleep', *ставам* 'get up', etc.
verb.change	18	Associative, Contextual Tasks	*залязвам* 'fade', *осветявам* 'brighten', *раста* 'grow', *ври* 'boil', *цъфтя* 'bloom', etc.
verb.weather	8	Thematic and Contextual Tasks	*блести* 'shine', *духа* 'blow', *ръси* 'dew', etc.
verb.perception	10	Contextual Tasks	*гледам* 'look', *помирисвам* 'smell', *слушам* 'listen', *чувам* 'hear', etc.

[a]The type of task in which the verb is used – Associative, Contextual, Thematic

motion and verbs of contact dominate, while in the thematic tasks the preference is given to verbs of weather, body, and consumption. The contextual tasks combine verbs from different thematic areas (e.g., verbs of consumption, verbs of change, verbs of contact, motion verbs).

For the analysis of verb associations, we rely on the associative knowledge of a frame and its frame elements. Thus, in the associative tasks the frame elements are offered as picture stimuli which were chosen as possible realisations of a frame element AGENT is a person or an animal, THEME is a physical object (a ship, a train, candle, a light bulb), INSTRUMENT is again a physical object (a shovel, scissors), LOCATION is a bench, etc.

The observations on verbs in context in the contextual tasks rely on sentences (and texts) within particular thematic areas. Context allows us to observe the frame elements – both core and non-core – in specific sentences.

4 Analysis and semantic description of basic vocabulary verbs

The analysis is based on three steps: 1) confirmation (or not) of the basic vocabulary hypothesis (using the percentage ratio); 2) observations on verb meaning competence using the semantic-syntactic structure of the target verbs; 3) observations on the statistical correlation of the results from the thematic and the association tasks.

In order to investigate the semantic structure of the target verbs and the respondents' knowledge about the verb, we use the information for semantic frames from FrameNet and lexical information for nouns from WordNet.

The semantic frames are schematic representations of situations and their participants – actors, circumstances, and other roles – which are elements of the frame. The frame elements have a name, a definition, and a semantic type, plus a specification of the relations between them (Fillmore 2000).

In FrameNet, all verbs sharing same semantic frame analyses evoke the same situation and share common frame elements and frame element relations (Ruppenhofer et al. 2016). For example, the verbs *hate, adore, despise, grieve, like, love* are analysed by the frame `Experience_focused_emotion` (Koeva 2021b).

We also investigate the semantic and syntactic compatibility between verbal predicates and the nouns that are realisations of their frame elements using the idea of conceptual frames as defined in (Koeva 2021a). They are abstract structures that determine the semantic and syntactic compatibility between verbal predicates and the concrete nouns evoking their frame elements. A particular

conceptual frame can be associated with a particular semantic class, and each frame element of the conceptual frame is associated with a set of nouns that are compatible with the predicate (Koeva 2021b).

Our observations on identification of frame-evoking lexical units in sentences lay on the hypothesis of conceptual frames that any of the noun synsets in WordNet can be a lexical representation of frame elements, thus the highest synset in a hyponym-hypernym tree can be chosen to describe the frame element (Koeva & Doychev 2022). For example, we may predict that {food} may evoke the frame element Food of the frame `Apply_Heat`. Thus, {food} and its hyponyms {meat}, {pasta}, {breakfast food}, {fish}, {seafood} can be analysed as elements of the frame evoked by the verb *boil*. Of course, not all synsets in a given tree would be appropriate for the specific frame element of a given frame.

Therefore, in the analysis of semantic classes of the target verbs below, we rely on the description of semantic frames, and comment on the pilot results with a view to the target verbs' semantic frames, their semantic and syntactic combinability with concrete lexical realisations of the frame elements.

The structure of a conceptual frame consists of three sections: lexical, morphological, and frame (syntactic and semantic) (Koeva & Doychev 2022). We will focus on the last section, where verbs are represented with semantic (as semantic restrictions to a given frame element, based on the WordNet noun synsets) and syntactic (as phrases – candidates for frame elements in Bulgarian) frame representation. As mentioned in (Koeva & Doychev 2022), semantic restrictions to a given element follow a certain context, and selective restrictions act between a predicate evoking a specific semantic frame, and its frame elements. Therefore, the two may differ for each case.

4.1 Analysis of thematically related verbs

The thematic tasks use the notion of a thematic area as defined with respect to the semantic class of the verbs. For example, the thematic area covering universal human needs and basic human activities such as *nutrition, body movement, health*, is defined by the usage of verbs related to bodily functions and form (*injure, lose weight, breathe*) or verbs of consumption (*drink, swallow, eat*). The target verbs are selected according to the WordNet semantic classes (for example, verb.body or verb.consumption), and the respondents had to choose five out of ten verbs according to a simple description of the thematic area. For example: (i) verbs representing actions or situations in the thematic area *weather* are extracted from the set of verbs labeled classified as verb.weather in WordNet; (ii) verbs for physical states or actions in the thematic area *human or animal body* are

extracted from the set of verbs classified as verb.body in WordNet; (iii) verbs in the thematic area of *using, buying or eating something* are extracted from the set ofverbs labeled with the semantic class verb.consumption in WordNet. Below we also comment on the verb groups which are part of the experiment with emphasis on their semantic description and representation.

4.1.1 Target verbs related to bodily functions and form

The general semantics of the thematic area can be illustrated by verbs, such as *контузвам* 'injure', *лекувам* 'cure', and *отслабвам* 'reduce (lose) weight', which correspond: to the frame `Experience_bodily_harm`: An Experiencer is involved in a bodily injury to a Body_part; to the frame `Cause_bodily_experience`: An Agent or Cause affects a Body_part of an Experiencer so that the quality of the Experiencer's perceptual experience of the Body_part is qualitatively changed; and to the frame `Cure`: a Healer is treating and curing an Affliction (injury, disease, or pain) of the Patient, whose Body_part is the specific area which is treated. The frame element Body_part is shared by all three frames.

The verb *дишам* 'breathe' with the semantic class verb.body is under discussion, as the body (or parts of it) is not explicitly realised frame element, but is implicit in the meaning of the verb 'for man and animal – I take air through my nose and mouth into my lungs and let it out after absorbing the oxygen necessary for the body' (see DBL).[4] The verb has the highest frequency of use in the Bulgarian National Corpus compared to the other four target verbs. The semantic frame that the verb evokes is `Breathing` with core frame elements Agent and Air.

The verb *поруменявам* 'blush' has low frequency, does not occur in selected textbooks and was not among the targeted verbs, but the results of the experiment show that it was recognised as a verb related to the human body. Verbs associated with the `Body_description_part` frame, typically evoke an `Individual` through a salient body part. The frame element Figure profiles the specific body part.

The results show that the score is higher if a verb's meaning is related to an experience directly affecting the body. The target verbs activate a physical dimension of the human body and its state, while the other (non-target) verbs evoke the person as a subject who can perform the action indicated by the verb, such as to sprinkle something on an object (*поръсвам* 'sprinkle' is verb.contact analysed by the frame `Filling`, with core frame elements Agent, Cause, Goal,

[4]Dictionary of the Bulgarian Language, https://ibl.bas.bg/rbe/

Theme); to eat some food (*хапвам* 'eat' is verb.consumption analysed by the frame `Ingestion`, with core frame elements Ingestor, Ingestibles); to understand something with one's mind (*разбирам* 'understand' is verb.communication analysed by the frame `Awareness`, with core frame elements Cognizer, Content, Expressor, Topic); to draw with my hand (*чертая* 'draw' is verb.contact analysed by the frame `Create_physical_artwork`, with core frame elements Creator, Representation); to fasten a garment on my body (*закопчавам* 'fasten' is verb.contact with frame `Closure` and core frame elements Agent, Container_portal, Containing_object, Fastener).

Table 2 shows the result: the third and the fourth column give the frame and the frame elements in FrameNet.[5]

Table 2: Results and the frames of the target verbs related to bodily functions and form

Verb	7-10 yrs	11-14 yrs	Frame	Frame elements
контузвам 'injure'	14.60%	16.80%	`Experience_bodily_harm`	Body_part; Experiencer
отслабвам 'lose weight'	16.00%	18.90%	`Cause_bodily_experience`	Agent; Body_part; Cause; Experiencer
лекувам 'cure'	10.00%	11.50%	`Cure`	Affliction; Body_part; Healer; Medication; Patient; Treatment
поруменявам 'blush'	11.90%	9.50%	`Body_description_part` (manually chosen)	Figure; Individual
дишам 'breathe'	17.80%	20.00%	`Breathing`	Agent; Air

[5]The second and third columns of the table show the results for 7-to 10-year-olds and 11- to 14-year-olds.

4.1.2 Target verbs of weather

For this task we used the thematic area related to the environment.

The verb *гърми* 'thunder', 'WN: be the case that thunder is being heard', and the verb *ръси* 'sprinkle' in the third person singular in the sense of 'DBL: rain, snow, etc. – fall, rain continuously' (DBL)[6] have low frequency both in the Bulgarian National Corpus and in the textbooks, but receive a high score among the answers, as can be seen in Table 3. The verb can be analysed using the frame `Make_noise`, with the core frame elements Noisy_event, Sound, and Sound_source.

The features shared by the verbs, labeled under the semantic class verb.weather, are environmental conditions – temperature, precipitation, wind, and sun as Source of weather – which activate a particular Place, Area, and Time, and respondents made an unambiguous choice of verbs related to weather.

Table 3: Results and the frames of the target verbs related to weather

Verb	7-10 yrs	11-14 yrs	Frame	Frame elements
вали 'rain'	22.40%	22.30%	`Precipitation`	Precipitation; Place; Time
духа 'blow'	22.40%	21.90%	`Motion`	Area; Direction; Distance; Goal; Path; Source; Theme
грее 'shine'	19.70%	20.50%	`Light_movement`	Emitter; Source; Beam; Path; Goal
гърми 'thunder'	21.80%	21.90%	`Motion_noise`	Area; Goal; Path; Source; Theme
ръси 'sprinkle'	11.60%	8.70%	`Precipitation`	Place; Precipitation; Time

[6]Dictionary Of The Bulgarian Language, https://ibl.bas.bg/rbe/

4.1.3 Target verbs of consumption

The results for the verbs in the semantic class verb.consumption (Table 4) confirm the assumption of the experiment that the respondents have acquired the basic vocabulary and have developed an active vocabulary covering different areas of human activity. The hypothesis that the selected verbs belong to the basic vocabulary, which contains lexical units that are used freely and spontaneously in everyday language practice, is confirmed. The high percentage of chosen verbs that belong to the same thematic area, suggests that the respondents can freely use the set of words and expressions, associated with the set of frame elements of the semantic frame `Ingestion` with core frame elements INGESTOR (person or animal that eats or drinks) and INGESTIBLES (objects that are ingested). In the case of the verb *дъвча* 'chew', the frame profiles the process rather than the consumption: in the frame `Grinding` a subject frame element GRINDER realises cause GRINDING CAUSE of an object PATIENT to be broken into smaller pieces.

Table 4: Results and FrameNet realisation of target verbs related to consumption

Verb	7-10	11-14	Frame	Frame elements
пия 'drink'	20.20%	20.70%	`Ingestion`	INGESTOR; INGESTIBLES
гълтам 'swallow'	19.20%	19.10%	`Ingestion`	INGESTOR; INGESTIBLES
хапвам 'eat'	19.20%	20.70%	`Ingestion`	INGESTOR; INGESTIBLES
дъвча 'chew'	19.20%	21.10%	`Grinding`	GRINDING_ CAUSE; GRINDER; PATIENT
гриза 'nibble'	19.20%	17.10%	`Ingestion`	INGESTOR; INGESTIBLES

4.1.4 Results of the experiment on thematically related verbs

As a result of the experiment, it was found that more than 80% of the participants (for both age groups) recognised the verbs of weather (*вали* 'rain' – 22.40%[7]/

[7]The percentages are given in the following order: 7- to 10-year-olds / 11- to 14-year-olds.

22.30%; *духа* 'blow' – 22.40% / 21.90%; *гърми* 'thunder' – 21.80% / 21.90%; *грее* 'shine' – 19.70% / 20.50%; *ръси* 'sprinkle' – 11.60% / 8.70%, against 0% and 2% for the other verbs *jump, sleep, dream, decide, carry*), as well as the verbs of consumption (eating and drinking) (*пия* 'drink' 20.20% / 20.70%; *хапвам* 'eat', *дъвча* 'chew', *гълтам* 'swallow', and *гриза* 'nibble' with the same result for the first age group 19.20% and for the second age group 20.70%, 21.10%, 19.10% and 17.10%). It is noted that the verb *ръси* 'sprinkle' means 'rain, snow, etc. – fall, rain heavily and continuously' is not clearly recognised as a weather-related verb, with a percentage score of 11.60% / 8.70%. Variations in choice are observed for verbs that activate semantic frames for actions related to bodily functions and form, with percentage distributions between 10% and 20% for the target verbs *контузвам* 'injure', *отслабвам* 'lose weight', *дишам* 'breathe', *поруменявам* 'blush', *лекувам* 'cure', and between 5% and 15% for the rest of the verbs: *sprinkle, bite, understand, draw* and *fasten*.

The results give us reason to consider the target verbs as part of the basic vocabulary of both age groups.

The Pearson correlation coefficient is used to measure the correlation of responses between the two groups. The number of responses, correlation and diagram can be seen in Figures 1, 2 and 3.

The result for the three groups of verbs: verbs related to bodily functions and form, verbs of weather, and verbs of consumption, is $R = 0.943702806$, $R = 0.994709117$, and $R = 0.993740859$, respectively.

This is a strong positive correlation, which means that high X variable scores go with high Y variable scores (and vice versa).

This gives us reason to conclude that high selection values in a given group increase or at least confirm the values of the other group. If a verb has a high percentage ratio in 11- to 14-year-olds, the verb may be part of the basic vocabulary for that group, but also it will be part of the basic vocabulary of the 7- to 11-year-olds, and vice versa. In other words – if a verb is mastered at the age of 7-10, then it is more likely to be part of a person's basic verb vocabulary at the age of 11-14. And vice versa – if a verb is part of the basic verb vocabulary at the age of 11-14, it will be part of the basic verb vocabulary at a younger age.

On the other hand, the correlation between the frequency of use of the target verbs and the responses for each age group is different. The coefficient shows a moderate positive correlation between verb's frequency of use and responses of verb of weather, and also between verb's frequency of use and responses in the 11-14 age group of verbs related to bodily functions and form, while in the other age group the correlation is defined as technically a positive one albeit weak. With verbs of consumption, the correlation also differs between the two age

	7 to 10 years	11 to 14 years
injure.v	32	16
lose weight.v	35	18
breathe.v	39	19
blush.v	26	9
cure.v	22	11
sprinkle.v	9	1
eat.v	33	12
understand.v	7	6
draw.v	7	1
fasten.v	9	2
correlation score: 0.943703		

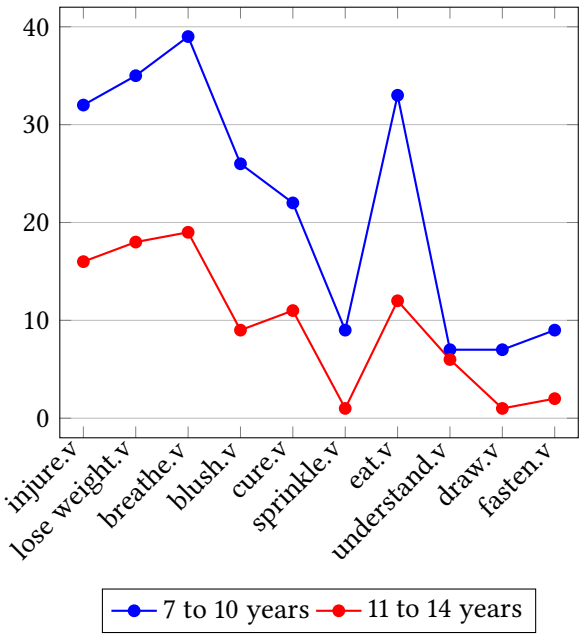

Figure 1: Correlation score of verbs related to bodily functions and form

	7 to 10 years	11 to 14 years
rain.v	33	112
blow.v	33	110
shine.v	29	103
sprinkle.v	17	44)
thunder.v	32	107
blush up.v	0	4
jump.v	1	6
sleep.v	1	6
dream.v	0	5
wear.v	1	6
correlation score: 0.994704		

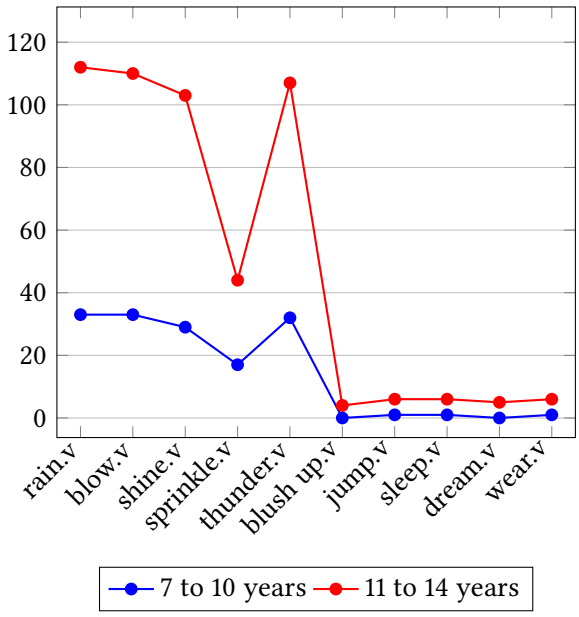

Figure 2: Correlation score of verbs of weather

	7 to 10 years	11 to 14 years
drink.v,	21	62
swallow.v,	20	57
eat.v,	20	62
chew.v,	20	63
nibble.v,	20	51
burn.v,	0	2
wash.v,	1	1
shave.v,	0	0
beat.v,	1	1
run.v	1	0
correlation score: 0.993741		

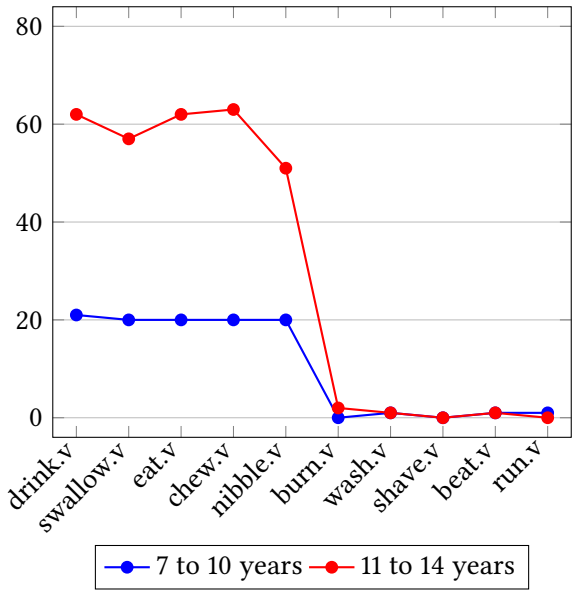

Figure 3: Correlation score of verbs of consumption

groups: the correlation in the 7-10 age group is technically a negative one, albeit weak, while the correlation in the age group 11-14 is again weak, but technically a positive one (see Table 5). Hence, the correlation between the frequency of use and the responses is not a significant factor in determining the dependence between the two variables.

Table 5: Correlation score between frequency and responses

verb.weather	frequency	7–10 years	11–14 years
rain.v	39.5	33	112
shine.v	22.29	29	103
blow.v	10.49	33	110
thunder.v	3.96	32	107
sprinkle.v	0.31	17	44
Correlation score		0.51693	0.55848

verb.body	frequency	7–10 years	11–14 years
breathe.v	51.61	39	19
cure.v	24.66	22	11
injure.v	22.1	32	16
lose weight.v	8.18	35	18
blush.v	4.38	26	9
Correlation score		0.48371	0.52761

verb.consumption	frequency	7–10 years	11–14 years
eat.v	323.88	20	62
drink.v	40.79	21	62
swallow.v	7.44	20	57
chew.v	4.6	20	63
nibble.v	2.12	20	51
Correlation score		-0.1401	0.38324

4.2 Analysis of target verbs via associative stimuli

Word associations are among the key storage mechanisms in recall (Glanzer 1972). Considering the intuitive nature of associations between the meaning of the verb, the semantic frame it evokes and its frame elements, we deem these types of tasks as the least difficult ones.

Association tasks are organised in two types.

4.2.1 Target verbs evoked by picture stimuli

In the first type, four verbs are associated with a given picture, where at least one of the verbs refers to the main sense and is assumed to be part of the basic vocabulary, without additional encoding of the manner of action (which may be encoded by prefixes, suffixes, etc.). Results confirm that respondents prefer verbs from the core vocabulary, where the picture stimulus represents an element of the respective frames (mostly core frame elements, but non-core ones as well).

For example, the picture stimulus of a running man is most often associated by respondents with the verbs *тичам* 'run' (51.2% among 7- to 10-year-olds; and 49.6% among 11- to 14-year-olds) and *бягам* 'run' (39.5% among 7- to 10-year-olds; and 44.7% among 11- to 14-year-olds). Both verbs belong to the synset {тичам; бягам} – {run}, verb.motion, with the definition 'WN: move fast by using one's feet, with one foot off the ground at any given time'. The frame is `Self_Motion`; the picture stimulus activates the core frame element – Self_Mover.

The picture stimulus of a shovel for the non-core frame element Instrument is associated with the verb *копая* 'dig' in 70% of responses, followed by *рина* 'shovel' in 19.2%, while the prefixed verbs *разривам* 'dig up' and *прекопавам* 'dig up' are chosen by less than 10% of the respondents. The picture stimulus of a bench for the core frame element Location is associated with the verb *седна* 'sit', which may instantiate the frame `Change_posture` (it is preferred to the prefixed verbs *поседна* 'sit down' and *приседна* 'sit down' and the manner verb *клекна* 'squat').

The verb *светя* 'light, glisten' is associated with the picture stimulus of a light bulb, illustrating either both the Source and Emission core elements of the `Emanating` frame, or the core elements of Figure and Light of the frame `Location_of_light`. The verb is preferred over *осветявам* 'illuminate', *светвам* 'light up', *блесвам* 'shine'.

Hesitancy among respondents, demonstrated by a more heterogeneous distribution of choices, is related to ambiguity or other possible association with the picture stimuli. For example, the picture stimulus of a dog standing upright on all

fours is most often associated with the derivative *изчаква* 'await' (40%) among 7- to 10-year-olds, less often with the non-derived verb *чака* 'wait' (24.4%), followed by the manner verb *дебна* 'lurk' (17.8%), while the 11- to 14-year-olds prefer the non-derivative *чака* 'wait' and *дебне* 'lurk' (both with 36%), possibly due to the association of the dog with aggression (the verb *дебна* 'lurk' can be described by the frame `Attack`). The picture corresponds to the PROTAGONIST core element of the `Waiting` frame.

Among all the tasks, clear preference (by over 50% of the respondents with clear margin with respect to the other possible choices) is given to basic (mostly non-derived) verbs referring to a simple action without additional specification of the manner of action, namely: *тичам* 'run', *бягам* 'run'; *копая* 'dig' (*рина* 'shovel', *светя* 'shine' (*осветявам* 'shine on', *прегръщам се* 'hug' (*гушкам се* 'cuddle', *седна* 'sit' (*поседна* 'sit', *подстриже* 'cut (hair)' (*оформя* 'form', *подрежа* 'cut', *налея* 'pour', *сипя* 'pour', *горя* 'burn' (*изгарям* 'burn out'.

4.2.1.1 Correlation scores of the experiment results

The correlation scores between the best and the second best answer (in %) are negative (as the children systematically have given a single answer) and statistically significant (Spearman's rho coefficient of the answers of 7- to 10-year-olds is $r_s = -0.74772$, and of the answers of 10- to 14-year-olds is $r_s = -0.82805$). The correlation scores between the best answers of the two age groups show strong positive correlation and are statistically significant ($r_s = 0.84242$) – see Figure 4.

4.2.2 Target verbs evoked by situation stimuli

In the second type of tasks, the picture stimulus also represents a possible realisation of a core frame element but respondents have to choose among ten verbs from different synsets, which may be analysed by different semantic frames. Five of the ten verbs are expected to be chosen, three are completely inappropriate, while the rest two are appropriate to a certain degree.

For example, the picture stimulus showing objects related to eating and drinking is associated with the verbs *закусвам* 'to eat breakfast' (20.9%[8] among 7- to 10-year-olds; and 21.9% among 11- to 14-year-olds), *хапвам* 'snack' (17.4% and 16.7%, respectively), *ям* 'eat' (17.4% and 16%, respectively), *похапвам* 'snack' (15.1% and 16%, respectively), and *пия* 'drink' (15.1% and 15%, respectively). All five verbs may be analysed by the frame `Ingestion`, with the picture stimulus corresponding to the core frame element INGESTIBLES. The preference for the manner verb

[8] All answers are equal to 100 percent, and the ratio is calculated accordingly.

Preferred verb	7 to 10 years (%)	11 to 14 years (%)
hear.v	73.8	79.3)
light.v	73.2	69)
dig.v	62.2	71.9)
embrace.v	82.9	86.8)
run.v	51.2	49.6)
burn.v	68.3	68.3)
await.v	24.4	36)
sit.v	80	61.5)
cut hair.v	53.6	54.8)
fill.v	44.4	44.6
Correlation score: 0.887412 (strong positive), r_s = 0.84242, p (2-tailed) = 0.00222		

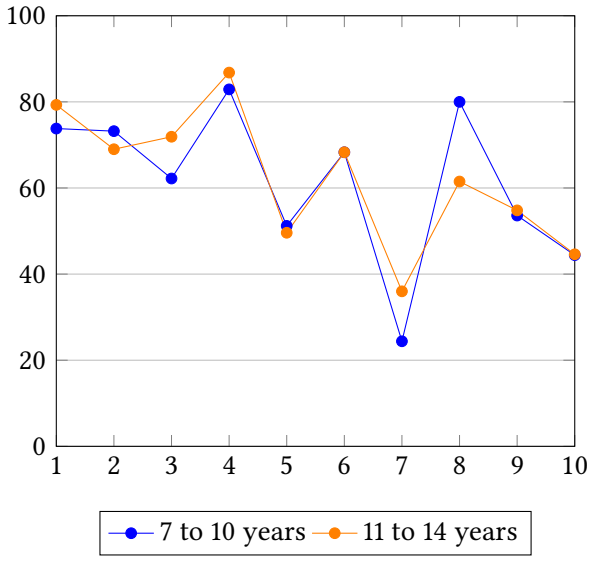

Figure 4: Correlation score between best answers of the two age groups(blue line – 7- to 10-year-olds; orange line – 10- to 14-year-olds).

закусвам 'eat breakfast' may be related to the time constraint in the sentence stimulus (*Сутрин* 'in the morning' which illustrates the non-core frame element Time). The picture stimulus – arranged objects related to food – also determines the next most frequent choices – *готвя* 'cook' (4.7% and 8.2%, respectively) and *подреждам* 'arrange' (3.5% and 3.9%, respectively).

Here, the disagreement among the focus groups is much lower probably due to the larger set of possible answers. The correlation scores between the highest probable and the highest less probable answer show weak positive correlation, but are not statistically significant (Spearman's rho coefficient of the answers of 7- to 10-year-olds is $r_s = -0.74772$, and of the answers of 10- to 14-year-olds is $r_s = -0.82805$).

4.2.2.1 Correlation scores of the experiment results

The correlation scores between the average of the target answers between the two age groups again show strong positive correlation and are statistically significant ($r_s = 0.81818$) – see Figure 5.

This basically confirms the conclusion from the thematically related tasks that there is correlation between basic verbs acquired by the two age groups. However, verbs' senses do not associate with each other.

Most of the verbs in both tasks evoke semantic frames that are linked to the `Event` top frame, as follows.

The frame `Motion` is linked via inheritance to `Self_motion`, `Motion_noise`, and `Motion_scenario` via usage to `Departing`, `Bringing`, `Removing`, `Emanating`, and via subframe to `Halt`.

The frame `Intentionally_affect` is linked to `Cutting`, `Education_teaching`, `Arranging`, `Grooming`, as well as to `Communication`, which, in its turn, is linked via inheritance to `Communication_manner` and `Communication_noise` and via usage to `Questioning`.

The frame `Intentionally_create` is linked to `Cooking_creation`, `Create_physical_artwork`, which is linked to `Create_representation`.

The frame `Intentionally_act` is linked to `Change_posture`, `Perception_active`, and `Manipulation`, which is further linked via inheritance to `Ingestion`.

The picture stimuli most often activate core frame elements such as: [Agent | Communicator | Speaker | Teacher | Creator | Protagonist | Self_mover (Sentient)] (where [Protagonist and Self_mover can also be represented by an animal); [Theme | Vehicle (Physical object)], as well as Goal, Body_part, Patient, Source, Ingestibles, Medium, Message. The non-core frame elements include: Instrument, Means, as well as Source, Goal, Body_part.

Average of the target verbs		
Subtask	7 to 10 years (%)	11 to 14 years (%)
1	17.68	17.68
2	15.94	15.94
3	16.66	17.56
4	16.14	16
5	17	17.64
6	18	17.64
7	13.86	13.06
8	18.22	17.36
9	18.42	18.26
10	17.18	17
Correlation score: 0.937487 (strong positive) r_s = 0.81818, p (2-tailed) = 0.00381		

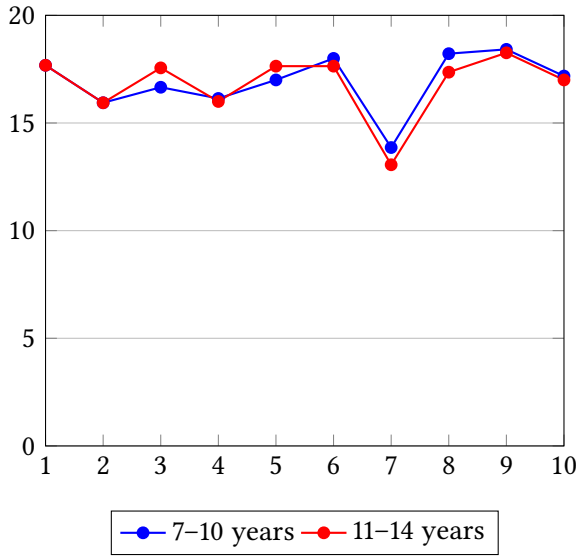

Figure 5: Correlation score between the recognised target verbs between the two age groups(blue line – 7- to 10-year-olds; orange line – 10- to 14-year-olds).

The most frequent non-core frame element, evoked by the picture stimuli, is Instrument, while the most frequent core frame element is Theme. The picture stimulus may activate the preference to one verb with a core frame element instead of another with a non-core frame element – for example, a picture of an ear is most often associated with the verb *чувам* 'hear' (74% and 79.3% in the two focus groups, respectively), which may be analysed by the frame `Perception_experience` with the core frame element Body_part; and less often with the verb *слушам* 'listen', which may be analysed by the frame `Perception_active` where the Body_part frame element is non-core (and unexpressed).

Hesitancy among respondents – demonstrated when a verb is chosen by a number of respondents below the mean – is observed if the picture stimulus does not entirely meet the characteristics of the frame element. For example, the picture of the dog standing, mentioned above, is associated with the verbs *чака* 'wait' by 36% among the 11- to 14-year-olds (but by 24.40% among the 11- to 14-year-olds, which is below the mean), while for the prefixed verb *изчаква* 'wait' the ratio is 37.80% among the 7- to 10-year-olds against only 8% of the 11- to 14-year-olds, while the preference to *дебне* 'lurk' is the opposite – 17.80% among the 7- to 10-year-olds (which is below the mean) and 36% among the 11- to 14-year-olds. The frame element Protagonist is specified as (Sentient), while the picture is of an animal. This is also the only task of the ten from the first type with such heterogeneous preference patterns (in five tasks, the preference is given to only one of the verbs; in two – to two verbs, while in the rest the two focus groups are unanimous about one verb, but do not agree on the second possible choice).

Table 6 illustrates the distribution of the respondents' choices in one of the tasks where the verbs may activate different semantic frames. The respondents have associated the meaning of the verbs stimulated by the meaning of a core frame element – for example, the verb *плува* 'swim' is analysed more often by the frame `Self_motion` and sentient Agent instead of the semantic frame Motion and Theme frame elements.

4.2.3 Prevailing semantic frame preferences in the respondents' selection

The results of the experiment on evoking verbs via associative stimuli allow us to summarise the most frequently selected frames of the experiment target verbs. The description of the frame includes a list of Bulgarian verbs, their semantic frame definition, their frame elements, information for the selectional specifics, and syntactic representation of the frame elements.

Table 6: Results of Association task – a picture of a ship (Activated frame element is underlined)

Verb	7-10 yrs	11-14 yrs	Frame	Frame elements
плава 'sail'	22.7%	20.3%	`Motion`	Area; Direction; Distance; Goal; Path; Source; Theme (PhysObj)
плува 'swim'	8%	7.5%	`Self_motion`	Area; Direction; Goal; Path; Self_mover (Sent); Source
акостира 'shore'	12.5%	11.4%	`Vehicle_landing`	Goal; Vehicle
потегля 'depart'	17%	18.3%	`Departing`	Source (Loc); Theme
спира 'halt'	9.1%	7.8%	`Halt`	Theme
превозва 'carry'	19.3%	19.9%	`Bringing`	Agent (Sent); Area; Carrier; Goal; Path; Source; Theme
износва 'carry off'	6.8%	5.9%	`Bringing`	Agent (Sent); Area; Carrier; Goal; Path; Source; Theme
продължава 'keep on'	4.5%	7.25%	`Activity_ongoing`	Agent (Sent); Activity; Duration
писука 'chirp'	0%	0.7%	N/A	N/A
мълчи 'be quiet'	0%	1%	N/A	N/A

1. Semantic Frame: `Cutting`

   Target verbs: *кълцам* 'chop'; *режа* 'cut' (*изрязвам, отрязвам, нарязвам, разрязвам* are language-specific verbs derived from the verb *режа* 'cut')

   Frame Definition: 'An Agent cuts an Item into Pieces using an Instrument (which may or may not be expressed).'

   Frame Elements: Agent // Item // Pieces

   Semantic and selectional specifics: NP Agent {person} *cut*.v NP Item {artifact}, PP Pieces {piece}

2. Semantic Frame: `Ingestion`

   Target verbs: *ям* 'eat', *похапвам* 'snack', *хапвам* 'eat', *закусвам* 'have breakfast', *гълтам* 'swallow', *гриза* 'nibble', *пия* 'drink'

   Frame Definition: 'An Ingestor consumes food or drink Ingestibles, and this entails putting the Ingestibles in the mouth for delivery to the digestive system.'

   Frame Elements: Ingestor // Ingestibles

   Semantic and selectional specifics: NP Ingestor {person} | {animal} *eat*.v NP Ingestibles {food} | {nutrient} | {meat} | {fish} | {vegetable} | {fruit}

   or

   NP Ingestor {person} | {animal} *drink*.v NP Ingestibles {beverage} | {alcoholic drink} | {water}

3. Semantic Frame: `Make_noise`

   Target verbs: *свиря* 'play', *викам* 'shout', *крещи* 'scream', *писука* 'chirp', *тананикам* 'hum'

   Frame Definition: 'A physical entity, construed as a point – Sound_source, emits a Sound. This includes animals and people making noise with their vocal tracts.'

   Frame Elements: Sound // Sound_source // Noisy_event

   Semantic and selectional specifics: *chirp*.v NP Sound {sound}, PP Sound_source {mouth}, PP Noisy_event {occurrence}

4. Semantic Frame: `Motion`

   Target verbs: *пътува* 'travel', *лети* 'fly', *духа* 'blow', *плава* 'float', *плува* 'drift' (English verbs for *плава* 'sail' and *плува* 'swim' are recognised as `Self_Motion` verbs in FrameNet.)

   Frame Definition: 'Some entity Theme starts out in one place (Source) and ends up in some other place Goal, having covered some space between the two Path. Alternatively, the Area or Direction in which the Theme moves or the Distance of the movement may be mentioned.'

   Frame Elements: Area // Direction // Distance // Theme // Source // Goal // Path

   Semantic and selectional specifics: PP Area, Direction, Distance, Path, Source, Goal can be in some case or another – {location} | {path} | {way} | {longness} | {land} | {area}, NP [Theme] {physical object1}

5. Semantic Frame: `Self_motion`

   Target verbs: *бягам* 'run', *тичам* 'run', *скачам* 'jump', *подскачам* 'jump', *подрипвам* 'caper'

   Frame Definition: 'The Self_mover, a living being, moves under its own direction along a Path. Alternatively or in addition to Path, an Area, Direction, Source, or Goal for the movement may be mentioned.'

   Frame Elements: Area; Direction; Goal; Path; Self mover; Source

   Semantic and selectional specifics: NP Self_mover {person} *run*.v PP Area, Direction, Goal, [Path], Source can be in some case or another – {location} | {path} | {way} | {land} | {area}

4.3 Analysis of contextually related verbs

The contextual competency of respondents was tested in two types of tasks with different level of difficulty – sentence usage of thematically related verbs and textual usage of verbs. The results of the two approaches are split. Respondents were able to handle thematically related verbs placed in a specific sentence context, but showed considerable difficulty in selecting verbs in a connected text.

4.3.1 Target verbs in a sentence context

The selected target verbs and the corresponding sentences denote situations and actions in respondents' everyday life. The results of the experiment showed that over 90% of the participants recognised the meaning of the verbs and use them correctly in the context. In this type of tasks we use the syntactic realisation of core frame elements of a semantic frame into the sentences as stimuli for the selection of target verbs. The sentences illustrating the verb's context represent situations evoked by semantic frames. The illustrative sentence for the verb *ям* 'eat' is analysed by the semantic frame `Ingestion`, *поръся* 'sprinkle' is analysed by the semantic frame `Filling`, *налея* 'pour' is analysed by the frame `Container_focused_placing`.

All sentences have core frame element Agent in the subject position with a null instantiation. The other core frame elements are: Ingestibles (*delicious and healthy food*); Theme (*sprinkle salt*) Goal (*on the toast*); Theme (*juice*); Goal (*into a large glass*).

In other sentences, however, the frames `Absorb_heat` *сваря* 'boil', `Apply_heat` *препека* 'toast', `Grinding` *настържа* 'grate' remain with null instantiations of the frame elements. For example, in the sentence *I boiled an egg* the Heat_source is not evoked, while in the sientence *Gonna grate some cheese*, the core frame element Static_object or Topic – the surfaces that rub against each other – *ренде* 'grater' is not instantiated. However, there are enough elements that are explicit, including for non-core frame elements, e.g., *I squeeze oranges for my favourite juice* [Goal] or *I will drink with relish* [Manner].

In the examples below we demonstrate the application of semantic-syntactic frames as stimuli in the concrete sentences from the tasks:

(1) *Сутрин обичам да* [pro]$_{\text{Ing}}$ ***ЯМ*** [`Ingestion`] [*вкусна и здравословна храна*]$_{\text{Ingbles}}$.

'In the morning I like to eat delicious and healthy food.'

(2) *Ето сега* [pro]$_{\text{Age}}$ ***ИЗЦЕЖДАМ*** [`Manipulation`] [*портокали*]$_{\text{Ent}}$ [*за любимия сок*]$_{\text{Result (non–core)}}$.

'Right now I am squeezing oranges for my favorite juice.'
* Unexpressed frame elements [Bodypart_of_agent] – the part of the Agent's body being used to manipulate the Entity.

(3) *Преди това* [pro]$_{\text{Age}}$ ***СВАРИХ*** [`Absorb_heat`] [*едно яйце*]$_{\text{Ent}}$ [*в тенджера*]$_{\text{Container}}$.

'Previously, I boiled an egg in a pot.'
* The source of heat treatment [HEAT_SOURCE] is not expressed.

(4) *В момента* [pro]$_{\text{AGE}}$ ***РЕЖА*** [Cutting] [*домата*]$_{\text{ITEM}}$ [*на парчета*]$_{\text{PIECES}}$.

'Currently, I am slicing the tomato into pieces.'

(5) *Взех филийки хляб, за да* [*ги*]$_{\text{FOOD}}$ [pro]$_{\text{COOK}}$***ПРЕПЕКА*** [Apply_heat] [*в тостера*]$_{\text{HEATINS}}$.

'I took slices of bread to toast them in the toaster.'
* Null instantiation frame elements are: [CONTAINER] – the object where food is stored and to which heat is applied, [HEATING_INSTRUMENT] – the object that emits heat, [TEMPERATURE_SETTING] – the temperature at which the food is processed.

(6) *Когато филийките са готови,* [pro]$_{\text{AGE}}$ *ще* [*ги*]$_{\text{GOAL}}$ ***НАМАЖА*** [Filling] [*с масло*]$_{\text{THM (}PhysObj\text{)}}$.

'When the slices are ready, I will butter them.'

(7) *След това* [pro]$_{\text{AGE}}$ *ще* ***ПОРЪСЯ*** [Filling] [*сол*]$_{\text{THM (}PhysObj\text{)}}$ [*върху филийката*]$_{\text{GOAL}}$.

'Then I will sprinkle salt on the slice.'

(8) [*Върху филийката*]$_{\text{PLACE (}non\text{–}core\text{)}}$ *ще* ***НАСТЪРЖА*** [Grinding] *и* [*малко кашкавал*]$_{\text{PAT}}$.

'On top of the slice, I will grate some cheese.'

(9) *Накрая* [pro]$_{\text{AGE}}$ *ще* ***НАЛЕЯ*** [Container_focused_placing] [*портокалов сок*]$_{\text{THM (}PhysObj\text{)}}$ [*в голяма чаша*]$_{\text{GOAL}}$...

'Finally I will pour orange juice into a large glass...'

(10) ... *и* [pro]$_{\text{ING}}$ *ще* ***ИЗПИЯ*** [Ingestion] [*с наслада*]$_{\text{MANNER (}non\text{–}core\text{)}}$ [*вкусната напитка*]$_{\text{INGBLES}}$.

'... and I will drink with delight the delicious drink.'

4.3.2 Target verbs in a text

The tasks aimed at textual usage of words gather information about the respondents' ability to acquire knowledge and to research, evaluate, and control this knowledge. Thus, they are the most difficult ones and combine a complex of

stimuli. The tasks imply that the participants have to take into account the lexical, grammatical, and morpho-semantic specificity of the verbs studied within a task (verbs with concrete and abstract meanings from all semantic classes, i.e., cognitive verbs, verbs of emotions, stative verbs, motion verbs, etc.).

Verbs that do not fit the specific usage in the text are also embedded in the sentences. They are selected on the following principle – a phonological competitor (a paronym) of the correct verb see (Examples 11d, 12b); a verb similar in meaning, but with a syntactic realisation incompatible with the context (Examples 12b, 13a) or which does not meet the lexico-grammatical requirements for the verb form – transitivity, reflexiveness etc. (Example 12c, 13c). These principles are illustrated in the short text part – adapted fragment of "Alice in Wonderland" used in one of the tasks.

(11) *Алиса* ***СКУЧАЕШЕ*** (страдаше (a), доскучаваше (b), нуждаеше (c)) *и си* ***МИСЛЕШЕ*** (приспиваше (d), успиваше (e), колебаеше (f)) *дали да* ***НАБЕРЕ*** (прибере (g), отнесе (h), обере (i)) *един букет от маргаритки в тежката следобедна горещина.*

'Alice was beginning to get very bored (suffer (a), beginning to suffer (b), need (c))[9] and she was **considering** (dozing off (d), starting to sleep (e), hesitating (f)) in her own mind whether **to pick** (take (g), bring (h), steal (i)) a branch of daisies in the hot afternoon.'

(12) *През това време един Бял Заек със светлочервени очи* ***ПОДСКОЧИ*** (посочи (a), поклати (b), изсмя (c)) *край нея.*

'At the same time a White Rabbit with pink eyes **ran** (pointed (a), shook (b), laughed (c)) close by her.'

(13) *Това не се* ***СТОРИ*** (оказа (a), престори (b), помисли (c)) *необикновено на Алиса и тя не* ***СЕ ИЗНЕНАДА*** (изстрада (d), измисли (e), сметна(f) *дори когато* ***ЧУ*** (слуша (g), попита (h), нахлу (i)) *как Заекът* ***СИ ГОВОРИ*** (въобразява (a), внушава (b), спори (c)) *„О, божичко, божичко!".*

'There **was** (did (a), pretended (b), thought (c)) nothing so very **remarkable** in that; nor **did** Alice **think** (suffer (d), invent (e), count (f) it so very much out of the way to **hear** (listen (g), ask (h), invade (i)) the Rabbit **say** (imagine (a), suggest (b), argue (c)) to itself, "Oh dear! Oh dear!"'

(14) *По-късно, като* ***РАЗМИСЛИ*** (замисли (a), измисли (b), сметна (c)), ***РЕШИ*** (разреши (d), представи (e), каза (f)), *че това е доста необичайно.*

[9]The translation of the alternative verbs in the tasks is literal and does not follow the above mentioned criteria – paronym, synonym, etc.

'When she **thought** (considered (a), imagined (b), reckoned (c)) it over afterwards, it **occurred** (allowed (d), presented (e), said (f)) to her that she ought to have wondered at this.'[10]

In addition, the task relies on respondents' knowledge of semantic and causative relations between sequences of verbs and uses those contextual relations as stimuli:

- semantic relation between target verbs as a contextual stimuli. For example: *скучаеше и си мислеше* (was bored and thought about); *размисли* (thought) ⟶ *реши* (decided) in Example 14;
- grammatical selection of a target verbs in the main clause and the target verb in the subordinate clause as contextual stimuli *скучаеше и си мислеше* (was bored and thought about) ⟶ *дали да набере* (whether to pick) in Example 11; *чу* (heard) ⟶ *как си говори* (her talk) in Example 12;
- grammatical combinability between target verbs prepositions and conjunctions (Examples 12, 13).

Another criterion of difficulty is the need of consideration of the information in the whole context used as last level stimuli of selection. All verb choices are presented shuffled below the text. Although in some sentences an alternative choice is possible (for example, Алиса мислеше и си каза дали да посочи... (Alice thought and said to herself whether she should indicate...) in composing the overall text, the alternatives are not acceptable.

As seen from Table 7, the largest difference in the responses between the two age groups – namely 7- to 10-year-olds and 11- to 14-year-olds – was observed at 35% and 65%, respectively, which can be explained by the complexity of the task. A total of 260 respondents filled in at least one verb position, with only 5% of respondents correctly filling in all positions in the task given, while 68% of responses were incomplete or incorrect, or possibly arbitrary.

Most errors were made with polysemous verbs, verbs of perception and verbs of cognition, such as *мисля* 'think', *изглеждам* 'look', or verbs like – *измъкна* 'pull', *свия* 'shrink', *вися* 'hang', as well as verbs with low frequency of use, such as *здрача се* 'dusk', *тъмнея* 'darken', some of which are on the periphery of the basic vocabulary list.

[10]The original Bulgarian translation of the text is adapted and simplified.

Table 7: Results of Contextual task – the adapted text part of *Alice in Wonderland*

Verb	Semantic class	7-10 yrs	11-14 yrs	Frame
чета 'read'	verb. cognition	32%	0.3%	`Reading_perception`: The READER attends to a TEXT to process its INFORMATION.
[*книгата*]$_{\text{TEXT}}$, *която* [*сестра ѝ*]$_{\text{RDR}}$ *четеше* '*the book her sister was reading*'				
набера 'pick'	verb. contact	38%	41.9%	`Food_gathering`: A GATHERER removes CROP ripe and ready to an accepted degree.
[*Алиса*]$_{\text{GATHERER}}$ *... да набере* [*един букет от маргаритки*]$_{\text{CROP}}$ (Example 11)				
подскоча 'jump'	verb. motion	11.2%	70%	`Self_motion`: The SELF_MOVER, a living being, moves under its own direction along a PATH.
През това време [*един Бял Заек*]$_{\text{SMOV}}$ *... подскочи* [*край нея*]$_{\text{PATH}}$ (Example 12)				
говоря 'speak'	verb. communication	1.7%	26%	`Statement`: the act of a SPEAKER to address a MESSAGE to some ADDRESSEE using language.
[*Заекът*]$_{\text{SPKR}}$ *да си говори* [*„О, божичко!“*]$_{\text{MSG}}$ (Example 13)				
свия се 'shrug'	verb. contact	11.9 %	27.6 %	`Posture`: An AGENT supports their BODY in a particular LOCATION.
... [*неговите уши*]$_{\text{BODP}}$ *се свиха* [*назад*]$_{\text{LOC}}$ '*his ears fell back*'; [*Зениците му*]$_{\text{BODP}}$ *се свиха* '*His pupils shrunk*'				
покажа се 'appear'	verb. change	6.7 %	1.8 %	`Cause_to_perceive`: An AGENT, causes a PHENOMENON to be perceived by a PERCEIVER
... [*в раззиналата му уста*]$_{\text{LOC}}$ *се показаха* [*белите му, остри зъби*]$_{\text{PHEN}}$ '*in his gaping mouth showed his white, sharp teeth*'				

The results for the four pre-established semantic frames activating the meanings of the verbs that were hypothesised to belong to the core lexicon and whose frame uniquely determines their position in the text, are tentative and correlated with the number of complete responses. This is probably also due to the unequal distribution of the number of responses obtained for the different options and in-between the different age groups.

5 Conclusion

In this article, we discussed the assumptions and subsequent results of a pilot survey aiming to explore whether the formulated language tasks can be used to test the respondents' acquisition of the semantics of a set of high frequency verbs, which are assumed to be part of the basic vocabulary. The language tasks activate the verbs' semantic frames and selectional preferences through different stimuli to explore the respondents' knowledge of specific semantic features – thematic verb groups, argument selection, syntagmatic usage of correct verb form.

Based on the results, we may conclude that the selectional preferences of the verbs we explored are mainly based on associations, and the choice of the respondents depends on the stimuli – whether pictorial, or textual. In addition, verbs that are considered part of the common lexis associated with a particular thematic area, are intuitively linked to a set of participants that are also part of the area.

The experiment helps us confirm or reject the hypothesis about the affiliation of the investigated verbs to the basic vocabulary. Respondents demonstrated a good understanding of the studied verbs' meanings related to nature, states, and actions from everyday life, as well as to material culture. They have internalised the usage of the verbs and their association with a semantic class.

The observations on the format of the experiment revealed that, in choosing the answers, respondents follow a "strategy" influenced by the uncontrolled online environment. The elective nature of the language games employed leads to their perception as a set of tasks that will ultimately be evaluated. Respondents tend to search for a single "most correct" answer, influenced by the presence of images and the arrangement of the verbs in the tasks.

The substantial difference in the results between the types of tasks is indicative in several respects such as: the difficulties in solving complex tasks, the role of the visual stimuli, the importance of knowing the selectional features of verbs such as reflexivity, transitivity, etc.

Despite the text being selected from a textbook for a lower grade, most of the respondents aged 7 to 10 faced considerable challenges with the last task. This

raises several questions related to their reading skills, as well as the readability of the text for respondents of a certain age, and the presentation method of the task in the survey.

Abbreviations

Age	Agent	Pat	Patient
BodP	Body_part	Phen	Phenomenon
Ent	Entity	Rdr	Reader
HeatIns	Heat_instrument	SMov	Self_mover
Ing	Ingestor	Spkr	Speaker
Ingbles	Ingestibles	Thm	Theme
Loc	Location		
Msg	Message		

Acknowledgements

This research is carried out as part of the project *Enriching Semantic Network WordNet with Conceptual Frames* funded by the Bulgarian National Science Fund, Grant Agreement No. KP-06-H50/1 from 2020.

References

Andonova, Elena. 2021. VERBA: adaptatsiya za balgarski ezik na vaprosnitsi za komunikativno razvitie [VERBA: Adaptation of questionnaires on communication development for Bulgarian]. In Svetla Koeva & Maksim Stamenov (eds.), *Dokladi ot Mezhdunarodnata godishna konferentsiya na Instituta za balgarski ezik Prof. Lyubomir Andreychin (Sofia, 2021) [Proceedings of the Annual International Conference of the Institute for Bulgarian Language Prof. Lyubomir Andreychin (Sofia, 2021)]*, 218–224. Prof. Marin Drinov Publishing House of the Bulgarian Academy of Sciences. DOI: 10.7546/ConfIBL2021.II.47.

Arnaud, Pierre & Henri Bejoint. 1992. *Vocabulary and applied linguistic*. Basingstoke: Macmillan.

Blanco-Carrión, Olga. 2006. Framenet as a corpus tool for the learning of second languages and for the lexical awareness of one's first language. *Porta Linguarum* 6. 67–76. http://hdl.handle.net/10481/30661.

Boyadzhiev, Todor. 2002. *Balgarska leksikologiya [Bulgarian lexicology]*. Sofia: Anubis.

Brysbaert, Marc & Andrew Biemiller. 2017. Test-based age-of-acquisition norms for 44 thousand English word meanings. *Behavior Research Method* 49. 1520–1523. DOI: 10.3758/s13428-016-0811-4.

Burov, Stoyan. 2000. *Balgarski ezik za chuzhdentsi: Osnoven kurs po leksika i gramatika [Bulgarian language for foreigners: Basic course in lexicon and grammar]*, vol. 1. Veliko Tarnovo: Faber.

Carter, Ronald. 1998. *Vocabulary: Applied linguistic perspectives*. London: Routledge.

Cohen, Andrew D. 1990. *Language learning*. New York: Newbury House.

Dalpanagioti, Thomai. 2021. A frame-inspired task-based approach to metaphor teaching. *Lexis – Journal in English Lexicology* 18. 1–22. DOI: 10.4000/lexis.5839.

Dimchev, Kiril. 2005. Otnosheniyata mezhdu ponyatiyata parvi ezik i vtori ezik v savremenniya nauchen diskurs [Relations between the concepts 'first language' and 'second language' in the modern research discourse]. In *Ezikovedski prinosi v chest na chlen-korespondent prof. mihail videnov [Linguistic contributions in honor of corr. member prof. mihail videnov]*. Veliko Tarnovo: Veliko Tarnovo University.

Dolgova, Natalia & Andrea Tyler. 2019. Applications of usage-based approaches to language teaching. In Xuesong Gao (ed.), *Second handbook of English language teaching*, 1–23. Cham: Springer. DOI: 10.1007/978-3-319-58542-0_49-1.

Duabeitia, Jon Andoni, Davide Crepaldi, Antje S. Meyer, Boris New, Christos Pliatsikas, Eva Smolka & Marc Brysbaert. 2018. MultiPic: A standardized set of 750 drawings with norms for six European languages. *Quarterly Journal of Experimental Psychology* 71. 808–816. DOI: 10.1080/17470218.2017.1310261.

Fellbaum, Christiane. 1990. English verbs as a semantic net. *International Journal of Lexicography* 3. 278–301. DOI: 10.1093/ijl/3.4.278.

Fillmore, Charles J. 1982. Frame semantics. In The Linguistic Society of Korea (ed.), *Linguistics in the morning calm*, 111–137. Seoul: Hanshin Publishing. http://brenocon.com/Fillmore%201982_2up.pdf.

Fillmore, Charles J. 2000. The FrameNet tagset for frame-semantic and syntactic coding of predicate-argument structure. In *Proceedings of the First Meeting of the North American Chapter of the Association for Computational Linguistics*. https://aclanthology.org/A00-2008.

Georgiev, Vladimir & Ivan Duridanov. 1959. *Ezikoznanie [Linguistics]*. Sofia: Nauka i izkustvo.

Glanzer, Murray. 1972. Storage mechanisms in recall. *Psychology of Learning and Motivation* 5. 129–193. DOI: 10.1016/S0079-7421(08)60441-7.

Jódar-Sánchez, José Antonio. 2018. FrameNet as a resource to teach Spanish as a foreign language. In María Luisa Carrió-Pastor (ed.), *Teaching language and teaching literature in virtual environments*, 121–149. Singapore: Springer. DOI: 10.1007/978-981-13-1358-5_7.

Karmiloff, Kyra & Annette Karmiloff-Smith. 2001. *Pathways to language: From fetus to adolescent*. Cambridge, MA: Harvard University Press. DOI: 10.2307/j.ctvjk2wz6.

Koeva, Svetla. 2021a. The Bulgarian WordNet: Structure and specific features. *Papers of the Bulgarian Academy of Sciences. Humanities and Social Sciences* 8(1). 47–70. https://www.papersofbas.eu/images/Papers_2021-1/Koeva_optimized.pdf.

Koeva, Svetla. 2021b. Towards Expanding WordNet with Conceptual Frames. In Sonja Bosch, Christiane Fellbaum, Marissa Griesel, Alexandre Rademaker & Piek Vossen (eds.), *Proceedings of the 11th Global Wordnet Conference*, 182–191. University of South Africa (UNISA): Global Wordnet Association. https://aclanthology.org/2021.gwc-1.21.

Koeva, Svetla & Emil Doychev. 2022. Ontology supported frame classification. In *Proceedings of the Fifth International Conference Computational Linguistics in Bulgaria (CLIB 2022)*, 203–214. Department of Computational Linguistics, IBL – BAS. https://aclanthology.org/2022.clib-1.23.

Koeva, Svetla, Ivelina Stoyanova, Svetlozara Leseva, Tsvetana Dimitrova, Rositsa Dekova & Ekaterina Tarpomanova. 2012. The Bulgarian national corpus: Theory and practice in corpus design. *Journal of Language Modelling* 0. 65–110. DOI: 10.15398/jlm.v0i1.33.

Meara, Paul. 1980. Vocabulary acquisition: A neglected aspect of language learning. *Language Teaching and Linguistics Abstracts* 13. 221–246.

Miller, George A., Richard Beckwith, Christiane Fellbaum, Derek Gross & Katherine J. Miller. 1990. Introduction to WordNet: An on-line lexical database. *International Journal of Lexicography* 3(4). 235–244.

Nisheva, Bozhana. 2013. Bezekvivalentnata leksika v uchebnitsite po balgarski ezik za chuzhdentsi [Lexical units with no equivalents in the textbooks of Bulgarian language for foreigners]. *OPERA SLAVICA* 23. 295–302. https://digilib.phil.muni.cz/sites/default/files/pdf/129036.pdf.

Popova, Velka. 2020. *Rechnik na detskite slovesni asotsiatsii [Children's dictionary of verbal associations]*. Shumen: Shumen University Press.

Ruppenhofer, Josef, Michael Ellsworth, Miriam R. Petruck, Christopher R. Johnson, Collin F. Baker & Jan Scheffczyk. 2016. *FrameNet II: Extended theory and practice*. Berkeley, California: International Computer Science Institute.

Schmitt, Norbert. 2007. Current perspectives on vocabulary teaching and learning. In Jim Cummins & Chris Davison (eds.), *International handbook of English language teaching* (Springer International Handbooks of Education 15). Springer, Boston, MA. DOI: 10.1007/978-0-387-46301-8_55.

Stenger, Irina & Tanya Avgustinova. 2021. Multilingual learnability and reaction time in online Slavic intercomprehension experiments. In *Dokladi ot Mezhdunarodnata godishna konferentsiya na Instituta za balgarski ezik Prof. Lyubomir Andreychin (Sofia, 2021) [Proceedings of the Annual International Conference of the Institute for Bulgarian Language Prof. Lyubomir Andreychin (Sofia, 2021)]*, vol. 2, 191–201. Prof. Marin Drinov Publishing House of the Bulgarian Academy of Sciences. DOI: 10.7546/ConfIBL2021.II.44.

Stoianova, Juliana. 2014. Littera et lingua. Series, dissertations. In http://slav.uni-sofia.bg/naum/liliseries/diss/2014/6.

Torrent, Tiago T., Ely Edison da Silva Matos, Frederico Belcavello, Marcelo Viridiano, Maucha Andrade Gamonal, Alexandre Diniz da Costa & Mateus Coutinho Marim. 2022. Representing context in FrameNet: A multidimensional, multimodal approach. *Frontiers in Psychology* 13. 1–20. DOI: 10.3389/fpsyg.2022.838441.

Vulchanova, Mila, David Saldaña & Giosué Baggio. 2020. Word structure and word processing in developmental disorders. In Vito Pirrelli, Ingo Plag & Wolfgang Dressler (eds.), *Word knowledge and word usage*, 680–708. Berlin: De Gruyter Moton. DOI: 10.1515/9783110440577-017.

Vulchanova, Mila & Valentin Vulchanov. 2021. Language development in typically developing children and children with developmental deficits: A concise qualitative review. *Papers of the Institute for Bulgarian Language Prof. Lyubomir Andreychin* XXXIV. 257–273. DOI: 10.7546/PIBL.XXXIV.21.10.