

# Cognitive Linguistics: A Neat Theory for Messy Data

Laura A. Janda and Stephen M. Dickey

*Abstract:* We outline some recent highlights in the application of cognitive linguistic theoretical and methodological approaches to the analysis of Slavic languages. A principal strength of cognitive linguistics is the way it focuses our attention on the continuous nature of linguistic phenomena. Rather than positing rigid categories and strict definitions, cognitive linguistics addresses the messy realities of language, facilitating the extraction of coherent patterns from the noise of human communication. We follow a thematic arrangement motivated by the types of variation we observe in language and the analyses proposed by Slavic linguists. These include variation across meaning and form, across modalities and genres, and across time and speakers.

## 1. Introduction: The Slavic World from the Perspective of Cognitive Linguistics

Roman Jakobson once stated that “the true difference between languages is not in what may or may not be expressed but in what must or must not be conveyed by the speakers” (Jakobson 1959/1971: 492). In other words, while it is possible to express the things that human beings want to communicate in all languages, languages differ widely in the categories that they require their speakers to pay consistent attention to. For example, speakers of Slavic languages face multiple grammatical requirements to mark gender, whereas speakers of Uralic languages can entirely ignore gender. Slavic languages are “satellite-framed” languages (Talmy 2000: Chapter 3), meaning that they tend to mark the path of motion in a satellite (a prefix) and as a result generally express the manner of motion in the root of a verb (as in Russian *vy-bežat'* [lit. out-run] ‘run out’). Romance languages like French are “verb-framed,” so they typically express the path of motion in the verb (as in *sortir* ‘exit’), and speakers often do not express the manner (though this is possible with an adverbial). Slobin (1996) has proposed that the grammatical categories that a language requires its speakers to mark consistently can have some effect on what those speakers pay attention to, and this “thinking for speaking” hypothesis (in

essence a milder form of the Sapir-Whorf hypothesis) has found support in numerous studies.

The same applies, *mutatis mutandis*, to linguistic theories. While one might ultimately succeed in getting a given point across regardless of theory, some theories are more amenable to given ideas and distributions of data, and some questions are facilitated to a greater or lesser degree by various theories. One of the most striking ways in which formalist and functionalist theories differ is in the discreteness of their approach to language. All scientific theories rely on metaphorical models as shortcuts to account for a much more complex reality. But all such models are by necessity flawed simply because they are models, and thus emphasize some aspects of reality while suppressing many others. Different models will bring different questions to the fore. Langacker (2006: 108–9, 146) states that “[a]n optimal metaphor does not guarantee an optimal description or constitute empirical verification. Conversely, a misleading metaphor does not prevent insightful description so long as one is fully aware of its limitations.” Whereas the metaphors of formalist theories are particularly discrete, often requiring absolute criteria, functionalist theories are more continuous, allowing for scalar phenomena.

Cognitive linguistics has generally focused on analyzing the semantic categories that constitute the meanings of linguistic units (whether individual morphemes or constructions of syntactically independent elements, see below) and determine their patterns of usage. In this respect, it views language as a dynamic emergent structure that is better characterized by patterns and relationships among items than by distinctions. One prominent way in which cognitive linguistics models relative continuity is by means of radial category structures in which a prototypical member of a category is related to relatively less prototypical members.

The focus of cognitive linguistics on continuity as opposed to crisp distinctions turned out to be historically propitious since cognitive linguistics blossomed at the same time that vast materials for studying linguistic variation and the tools for analyzing such variation went through radical changes. By “vast materials,” we refer to linguistic corpora and the electronic means to collect, share, and manipulate such authentic language data. Corpora give evidence not only of suspected language variation (often mentioned but not substantiated in reference works) but almost always reveal variation also where linguists have traditionally thought there were “rules”. When the big data of the information age became a reality, the discipline of statistics rose to the challenge, making great strides both in terms of models and access to powerful software. Cognitive linguistics has played a special role as a theoretical framework that accounts for the variation found in linguistic data and supports the interpretation of statistical outcomes. This volume contains an article on “Slavic corpus and computational linguistics” by Divjak, Sharoff and Erjavec (171–199), so we do not detail this type of investigation here,

but wish to emphasize that much of the relevant work cited therein has cognitive linguistic underpinnings.

Another continuous feature of cognitive linguistics is the fact that it does not propose an atomistic approach to the units of language. In other words, cognitive linguistics is not reductionistic, it does not assume that there are minimal units of language that can be assembled via composition into larger units. The basic unit of language is a construction (Goldberg 2006), a “symbolic assembly” (Langacker 2013) consisting of a phonological pole and a semantic pole, but constructions can be of varying levels of complexity, from very small (word-internal) to large (discourse level). This is the fundamental structure of “construction grammar”, described by Mirjam Fried in her article in this volume (241–276). Construction grammar is an organic part of cognitive linguistics, and although we mention some relevant works in passing, we leave the details to Fried.

Some overlap with these and other articles in this volume is inevitable, but we focus primarily on topics not addressed by other authors in this silver issue of *JSL*, specifically on the effects of variation in the Slavic languages and how they have been analyzed by means of cognitive linguistics. Within the limits of this overview article we cannot do justice to the full range of works that have represented cognitive linguistic analyses of Slavic material over the past fifteen years; our aim is merely to present some highlights, focusing on variation in Slavic data and how it is accounted for in cognitive linguistics. This survey represents both works that are explicitly framed in terms of cognitive linguistics as well as those that are highly compatible with cognitive linguistics.

Variation can be observed over many domains. There can be variation in the signal of language itself, in both the meaning and the form. Section 2 is devoted to variation in the signal. Variation can manifest itself across the ways in which language is used, in the genres of both spoken and written discourse, and in the use of co-speech gesture. Variation in genre and modality are addressed in Section 3. Finally language can also vary across space and time, through dialectal and sociolectal differences, and the evolutive processes of language acquisition and language change. These variations across speakers and times are the topic of Section 4.

## 2. Variation across Meaning and Form

As mentioned above, Langacker (2013) describes the symbolic assembly joining meaning and form as the fundamental structure of language, and this structure is not restricted to minimal units but relevant (at various levels of schematicity) at all levels of complexity. The symbolic assembly does not assume a one-to-one relationship between meaning and form. Instead, both

meaning and form may be complex and variable, and even the relationship between the two can vary along a scale from relatively autonomous to iconic.

The nature of meaning plays a central role in cognitive linguistics, which can be opposed to structuralist and formalist approaches in that it focuses on the internal structure of meaning rather than on boundaries or supposedly invariant “features.” The meaning of a given linguistic expression, category, or other meaningful structure is centered around a prototype, and extension relationships (often motivated by metaphor and metonymy) link to more peripheral members, forming a radial category. Membership in a radial category is scalar, as are the relationships among members. Radial categories have been used extensively in cognitive semantic analyses of a range of Slavic lexical and grammatical categories. In what follows we survey recent work on case and aspect, two of the most distinctive Slavic grammatical categories.

Cognitive linguistic analyses of case have employed radial categories to organize and make sense of the various functions of cases in Slavic languages. The posthumously collected works of Rudzka-Ostyn (Tabakowska 2000) trace some of the earliest analyses of this kind, focusing on the grammatical cases of Polish. Janda’s (1993) study on the meanings of the instrumental and dative cases in Czech and Russian took the indirect object use of the dative case as prototypical, with extensions entailing the absence of an accusative direct object (governed dative) and the absence of an agent (impersonal dative). The interpretation of the impersonal dative in Russian was further detailed by Fortuin (2000), and the Polish dative was examined by Dąbrowska (1997). This approach eventually expanded into full-scale analyses of all cases in both bare and prepositional uses in Czech and Russian (Janda 2002), textbooks for learners (Janda and Clancy 2002, 2006), and typological comparisons of case usage across Slavic which proposed a semantic map of case semantics for Russian, Polish, and Czech (Clancy 2006). Šarić (2008) focused on spatial concepts expressed by the prepositions *na*, *u*, *kod*, and *pri*, as well as the directional bare dative case in Bosnian/Croatian/Serbian (B/C/S). Janda’s and Clancy’s case textbooks for Russian and Czech for learners demonstrate the value of a cognitive semantic approach in instructional materials.

Aspect also became the subject of great attention in cognitive linguistics. Again, the functional orientation of cognitive linguistics combined with its emphasis on complex category structure resulted in new approaches to the category and its difficulties. Cognitive linguistics filled a gap in scholarship not only by developing approaches to the usage of the aspects but also by means of a renewed interest in the semantic nature of its major markers—verbal prefixes. Aspect and the prefixes are considered in turn.

Dickey (2000) applied basic principles of cognitive semantic analysis in a study of the variation in Slavic aspectual systems, organizing and analyzing the variation as the consequence of two conceptually contiguous categories, totality and temporal definiteness (i.e., uniqueness in time, cf. Leinonen

[1982]). On the basis of a variety of usage patterns and constructions (habituals, statements of fact, the historical present, running instructions,<sup>1</sup> coincidence with the moment of speech, ingressivity, and verbal nouns), this study established a broad east-west division in Slavic aspect. Accordingly Slavic aspectual systems break down into a western group (Czech, Slovak, Sorbian, and Slovene) motivated by a prototype of totality (comparable to the count/mass distinction for nouns), an eastern group (Russian, Ukrainian, Belarusian, and Bulgarian) motivated by a prototype of temporal definiteness (involving referential uniqueness and thus comparable to definiteness in nouns), and two transitional zones (B/C/S and Polish) that share both types of conceptualization.

Combining radial-category analysis with quantitative methods offers great potential for the comparative analysis of closely related Slavic systems in general. For example, Christensen (2011) combines these approaches to uncover subtle differences between the perfectivizing function of *po-* in Russian and Polish. Christensen shows that the semantic prototype of Russian *po-* is delimitativity, whereas the semantic prototype of Polish *po-* is distributivity, and that delimitative *po-* is relatively more productive in Russian than in Polish, and conversely, distributive *po-* is more relatively more productive << **'more' repeated intentionally?**>> in Polish than in Russian. Christensen's data and conclusions help us to understand how Polish serves as an aspectual transitional zone in Slavic. Another application of these methods is Moulton (2015), which examines subtle differences between reflexive verbs in B/C/S and Russian to show that the prototype of B/C/S *se* 'self' is true reflexivity, whereas the prototype of the Russian suffix *-sja* is middle voice. Moulton's data and conclusions are important for typological studies such as Kemmer (2013), which takes Russian as representative of Slavic.

The principles of cognitive linguistics have allowed linguists to recognize the parallels between the referential properties of nouns and verbs. Taking up the parallels between nouns and verbs in terms of bounding outlined in Langacker (1987) and applied to Russian by Mehlig (1996), Janda (2004) explores the metaphorical identification of perfective with the characteristics of solid objects (which are clearly bounded with definite shapes) vs. imperfective with fluid substances (which lack clear boundaries and inherent shapes). In Russian, for example, this metaphorical characterization is extended to domains of discourse (where imperfective "substances" form the background for perfective figures, or "objects") and pragmatics (where in certain circumstances imperfective imperatives are more polite, like soft substances, than perfective imperatives, like hard objects). The most recent development in this line of investigation is Janda and colleagues' (2013) hypothesis that Russian prefixes

<sup>1</sup> Instructions in the present-tense indicative that are given as a characterization of a process to be performed.

are verbal classifiers on a par with numeral classifiers in Asian <<Caucasian??>> and other languages (see below).

Over the last couple of decades, Slavic cognitive linguistics has been at the forefront in a revival in the study of Slavic verbal prefixes. This can be traced back to Janda (1986), which interpreted the meanings of four Russian verbal prefixes that shared the meaning of excess (*pere-*, *za-*, *do-*, and *ot-*) as coherent networks. For example, the prototypical meaning of *za-* involves a landmark that “designates the normal or canonical environment from which the trajectory deviates” (Janda 1986: 79). Depending on the environment, this meaning expresses deflection, entering a fixed state, excess, or inchoative action, and further extensions involve changes in the dimensionality of the landmark and the parameters of the trajectory. The ensuing interest in the functions of Slavic prefixes occurred naturally in cognitive linguistics, which recognizes that the division between grammar and lexicon is basically arbitrary and in any case a fluid one. This has allowed it to develop coherent analyses of the interacting lexical and grammatical functions of Slavic prefixes, which are the building blocks of verbal reference in Slavic languages.

Janda (1986) was followed up by a number of studies of prefixes in Slavic languages. Shull (2003) combined cognitive semantics and video experiments to analyze differences between the Russian and Czech systems of prefixation. Dickey and Hutcheson (2003) and Dickey (2005, 2007) have analyzed the function of *po-* and *s-/z-*. Šarić (2014) presents cognitive analyses of prepositions and prefixes in B/C/S. Kuznetsova (2015: Chapter 4) has elaborated on the notion of prototypicality with an analysis of the Russian prefix *pri-*, which has three submeanings (arrive, attach, and attenuate) that compete for recognition as the prototype. Ultimately Janda and colleagues returned to the issue with radial category analyses of all Russian prefixes that can be used to form Natural Perfectives (of the type *na-pisat'* ‘write’, *s-varit'* ‘cook’). They propose that aspectual prefixes in Russian, and in Slavic in general, behave as a verb classifier system analogous to numeral classifiers, which likewise show polysemy describable in terms of radial categories (Janda et al. 2013). Dickey and Janda (2015) have taken this line of research one step further by detailing typological parallels between numeral classifiers and all types of prefixed perfective verbs across the Slavic language family.

Slavic lexical items and semantic fields have been analyzed from a cognitive linguistics perspective. Rakhilina has made numerous contributions with analyses of the metaphorical behaviors of groups of lexemes in Russian, for example adjectives of size, temperature, and color; body parts, body positions; manner of motion; and bodily experiences such as pain (Rakhilina 2000; Majsak and Rakhilina 2007). For example, we learn from Rakhilina that the colors *zelenyj* ‘green’ and *želtyj* ‘yellow’ share negative associations due to their opposition to other more “human” colors, and that position verbs like *stojat'* ‘stand’ and *ležat'* ‘lie’ are applied to inanimate objects according

to their function rather than their orientation. In Russian, *posuda stoit na stole* 'the dishes stand on the table' even though they are flat horizontal objects because they are in use, whereas *posuda ležit v bufete* 'the dishes lie in the cabinet' when they are not in use. In both cases, the metaphorical source domain is human beings who are normally neither green nor yellow and who stand up when active but lie down when inactive. In this way Rakhilina and colleagues have fleshed out numerous metaphors that make up the "*rususkaja jazykovaja kartina mira*" (see also Šmelev 2002 and Zaliznjak et al. 2005). Similar studies have also been undertaken for other Slavic languages, most notably Będkowska-Kopczyk (2004) has investigated negative emotions in Slovene.

Near-synonymy is another topic that involves the structure of meaning, and cognitive linguists have concerned themselves with questions about how various factors influence the choice among synonyms. Janda and Solovyev (2009) analyzed 500 corpus sentences for each of six Russian synonyms for 'sadness' (*grust', melanxolija, pečal', toska, unynie, xandra*) and five synonyms for 'happiness' (*likovanie, naslaždenie, radost', udovol'stoie, vostorg*) according to their "constructional profiles," namely the statistical distribution of case marking on the noun and the presence (or absence) of prepositions. While there were some overall similarities (for all eleven nouns in question, the same five [preposition] + case constructions were most characteristic: *v* + Acc, *v* + Loc, bare Inst, *s* + Inst, *ot* + Gen), they showed that each synonym had a unique constructional profile. For example, *grust'* is most strongly attracted to the *s* + Inst construction, whereas *unynie* is most attracted to the *v* + Acc construction.

In an ambitious study, Divjak (2010) investigated the relationships among nine Russian "tentative" verbs meaning 'try': *probovat', pytat'sja, starat'sja, sil-it'sja, norovit', poryvat'sja, tščit'sja, pyžit'sja, tužit'sja*) by means of behavioral profiles. She analyzed 1,585 sentences from corpus data containing tentative verbs in finite form construction with an infinitive for 87 different parameters, among them: the tense, mood, and aspect of both the tentative verb and the infinitive; features of the clause, sentence, and subject of the verb, presence of adverbs, particles, negation, conjunctions; and semantic properties based on Apresjan's (1995) primitives of human behavior. Divjak's statistical analysis showed that it is possible to discriminate among near synonyms, and that the strongest predictors for the choice of one synonym over another were the tense, aspect, and mood marking on the finite verb and infinitive, followed by semantic properties of the subject and infinitive event. These findings, which were corroborated by experimental evidence, are at variance with both the lexicographic and psycholinguistic traditions and suggest that native speakers use distributional cues to develop a prototype, and also that groups of semantically similar words are clustered in the minds of speakers. Divjak (2010: 228) concludes that her study "shows how rich purely linguistic input really is, and how powerful the form-meaning relationship is."

Variation in form has been studied from the perspectives of phonology, morphology, and syntax. While in general phonology has been underrepresented in cognitive linguistics, Nessel (2008) has gone a long way to fill this void with a theoretical account of morphophonological alternations in the Russian verb stem. Nessel introduces “second-order schemas” as a means for modeling relationships between surface forms without recourse to underlying representations or ad-hoc rules. These schemas make it possible to model the relationships among paradigm forms such as: the relationship between the 3pl present and present active participle forms and the relationship between past tense and infinitive. Nessel ultimately argues that form itself has meaning, that for example the truncation and softening alternations in Russian are markers of non-past meaning.

Schematicity is likewise a central concept in Fidler’s (2014) groundbreaking work on onomatopoeia in Czech. Based on data from corpora, literature, comic strips, internet, and dictionaries, Fidler examines the intricacies of the form-meaning relationship at the level of the sounds themselves and their positions (onset, coda). She challenges the traditional assumption that sound and meaning are independent, suggesting that arbitrariness resides instead in the ways in which individual languages connect the two. Any given sound can be dissected in a potentially infinite number of ways. For example, the onset labial stop *p*- can be viewed as a complex articulatory gesture with many components, including “the process of building up the airstream behind the obstruction before rupture, the incapacity of the obstruction to hold the buildup, the instantaneity of the release of the air, the power of the released air, the speed of the air” (Fidler 2014: 178). Czech happens to focus on the building up of pressure in this gesture and thus associates it with bursting sounds and motions, such as *pif*, *paf* (the sound of bullets being fired), *plask*, *plesk* (the sound of hard objects hitting water), and *prásk* (the sound of something breaking apart upon impact). This particular phonosemantic association is just one option; other languages can focus on other parts of the articulatory gesture. In this way, the arbitrariness of onomatopoeia parallels the arbitrariness of semantic mapping relations like metaphor and metonymy in the naming of objects. For example, the word for ‘octopus’ exploits various possible construals across languages: in Russian it is *osminog* [lit. eight-legs] invoking an iconic metaphor between the tentacles and legs; in Czech it is *chobotnice* [lit. elephant-trunk-thing] invoking a different iconic metaphor between the tentacles and an elephant’s trunk; in Norwegian it is *bleksprut* [lit. ink-squirt] invoking a metonymic reference to the animal’s defensive behavior. In a similar fashion, Fidler (2014: 54) argues that “arbitrariness ... should be sought in the manner in which speakers choose some of all the aspects of each articulatory gesture as the most prominent. ... Thus... onomatopoeia... results from the cognitive process that operates also in other parts of language.” Fidler argues that the relationship of sounds to meaning in onomatopoeia extends beyond



iconic metaphors based on articulatory gestures, reaching also into grammatical categories such as aspect and serving as discourse markers.

The distribution of forms in corpora is meaningful as well, a fact accounted for in various studies based on behavioral profiling. Janda and Lyashevskaya (2011) found that the distribution of Russian verb forms according to subparadigms (nonpast, past, infinitive, imperative) was distinct for perfective as opposed to imperfective aspect and that this distinction was independent of the aspectual morphological markers (prefixes vs. suffixes). Kuznetsova (2015) challenged the Maslov (1984) criterion for identifying Russian aspectual pairs by substitution of partner verbs in the same grammatical constructions using corpus data and demonstrated that supposedly “paired” verbs vary greatly in the number and frequency of grammatical constructions that they share. Eckhoff and Janda (2014) used variation in distribution of Old Church Slavonic verb forms to show that it is indeed possible to sort verbs according to aspect already in the earliest Slavic texts, suggesting an early provenience for perfective vs. imperfective verbs.

The term “rival forms” applies to sets of morphemes and more complex constructions that express the same or very similar meanings, such as, for example: the use of the “theme-object” construction *gruzit' seno na telegu* ‘load hay onto the cart’ vs. the “goal-object” goal object construction *gruzit' telegu senom* ‘load the cart with hay’ in Russian; or in the alternative Czech expressions for ‘through the window’ *skrz okno* vs. *oknem*. Studies of rival forms discover the distribution of the competing forms and the strength of factors associated with the variant forms and furthermore try to reveal diachronic trends and meaningful motivations for the variation in form. For example, by far the strongest factors relevant for the distribution of Russian *gruzit'* ‘load’ constructions are the contrast among the imperfective and its three Natural Perfective partner verbs *nagruzit'*, *zagrutzit'*, and *pogrutzit'*, with lesser roles played by the use of active vs. passive voice (with a participle), and the naming of both the theme and the object vs. use of a reduced construction with only the theme or only the object. In terms of meaning, the distribution shows an association of the prefix *na-* with the goal-object construction, which makes sense given that the prefix refers to surfaces ‘on’, the use of *po-* almost exclusively with the theme-object construction, which makes sense since it refers to a change of state for the theme, and mixed use for *za-* which has many metaphorical uses (Sokolova et al. 2012).

Sergey Say has undertaken a number of studies on the rivalries among Russian syntactic constructions. For example, in Say (2013), the topic is the use of the dative case to mark experiencers as opposed to null marking or other markings (such as *dlja* ‘for’). Say discovers a hierarchy for the co-occurrence of dative NPs denoting sentient participants such that predicatives are prototypical, and there is decreasing compatibility with related short-form adjectives and even less with long-form adjectives. Based on analysis of corpus data, Say

identifies three basic types of predicatives: 1) the *xolodno* 'X feels cold' type for which the predicative's dative argument is not related to any of the adjective's arguments (the adjective has a meaning that does not presuppose an experiencer), 2) the *grustno* 'X feels sad' type for which the predicative's dative argument is related to the adjective's internal argument (the adverb does co-occur with the dative, but the adjective does not), and 3) the *prijatno* 'X feels it is pleasant' type for which the predicative's dative argument is related to the adjective's dative argument. Essentially Say puts forward a product-oriented schema according to which the constructional pattern is "blind" to the argument structure of the corresponding adjective, and constructional meaning prevails, either demoting or coercing the experiencer to the dative slot.

The study of patterns of variation in both meaning and form yield opportunities to find coherent regularities that otherwise might go unnoticed. For example, allomorphy is traditionally invoked only when the meaning of two (or more) forms is "identical," the forms are related (usually sharing the same etymology), and they are complementarily distributed. This strict definition excludes many significant relationships among linguistic expressions that are worthy of attention but involve meanings that are extremely close (but not identical) and/or forms that are unrelated. Dickey and Janda (2009) discovered a robust relationship between the semelfactive uses of the Russian *-nu* suffix and *s-* prefix as in *čixnut'* 'sneeze once' and *sglupit'* 'do one stupid thing.' It turns out that the distribution of these morphemes is largely, but not entirely, predictable based on the morphological classes of the verbs they attach to, with *-nu* strongly preferred for verbs suffixed in *-aj* (*zevat'* > *zevnut'* 'yawn > once'), *-\*š* (*svistet'* > *svistnut'* 'whistle > once') and representing non-productive I conjugation types (*lizat'* > *liznut'* 'lick > once') vs. *s-* strongly preferred for verbs suffixed in *-ova* (*malodušestvoovat'* > *smalodušestvoovat'* 'be cowardly > once'), *-i* (*grubit'* > *sgrubit'* 'be rude > once'), and *-\*ěj* (*robet'* > *srobet'* 'be shy > once'). Since the distribution is not complementary (there is some overlap), it would not be possible to invoke allomorphy in the strict sense; however, the behavior of these two morphemes closely approaches allomorphy and deserves notice. The study of such non-standard allomorphy has been further extended to other Russian prefixes in two dissertations. Makarova (2014) explores the attenuative/diminutive meanings of Russian *pri-* and *pod-* as cases of semantic overlap at the peripheries of their radial categories. Endresen (2014) investigates the range of possibilities, from prototypical allomorphy in the case of *raz-* vs. *ras-* and standard allomorphy in the case of *raz-* vs. *razo-*, through non-standard allomorphy in the cases of *s-* vs. *so-*, *o-* vs. *ob-* vs. *obo-*, *pere-* vs. *pre-*, *vz-* vs. *voz-*, *vy-* vs. *iz-*, to cases of non-allomorphy with *o(b)-* vs. *u-* and *pre-* vs. *pri-* vs. *pred-*. Together, these works show that traditional allomorphy only skims the surface of a host of significant relationships among forms that share meanings in language.

### 3. Variation across Modalities and Genres

Virtually all of the studies described in the previous section focused solely or primarily on the written modality, usually represented in a corpus. However, written language is of course secondary to spoken language, and spoken language is typically accompanied by gesture. This section briefly represents the spoken and gestural modalities in turn.

A major landmark in investigation of spoken Russian is the corpus and analysis offered in Kibrik and Podlesskaja 2009. They collected children's narrations of their dreams as sound files and transcribed them to include annotation for prosodic features, such as tone, accent, loudness, tempo, and pauses. A major goal of their research was to understand how spoken discourse is segmented, since of course it lacks punctuation that signals segmentation in written texts. They found that the basic unit of spoken Russian is an "elementary discourse unit" (EDU) that can be identified on the basis of prosody. The prototypical EDU contains 2–5 words and corresponds to a clause. However, over 30% of EDUs deviate from this prototype, and most of the non-prototypical EDUs are shorter. Some subclausal EDUs serve discourse functions (consisting of discourse markers, or undergoing truncation when the speaker feels a need to restart); however, most of these "small" EDUs have other functions, namely they repeat information, elaborate on it, or simply divide up a complex clause into more convenient portions. It is precisely the subclausal EDUs that escape notice in linguistic analyses of the written modality, where they are underrepresented or altogether absent. Ultimately Kibrik and Podlesskaja find that the sentence is harder to define in spoken discourse, since it depends also upon the prosodic habits of individual speakers, but that the sentence is also less relevant, since it serves only at an intermediate level of organization, between the EDU and discourse episode.

While relatively little has been published to date on the use of co-speech gestures in Slavic (some exceptions are a dictionary of Russian gestures by Grigor'eva et al. 2001 and studies of how various concepts are conveyed by gesture in Polish: Antas 2001 and Załazińska 2001, promising research is underway, particularly in Moscow. Kibrik (2010) undertook an experiment comparing the amount of information conveyed by three different modalities: segmental (verbal) signals, prosodic signals, and visual signals (gesture and other visible cues). The experimental stimulus was a 3.3 minute videoclip from a television serial, preceded by an 8 minute clip from the same serial to set the context. Participants were asked to answer questions based on the 3.3 minute clip and were divided into eight groups that received: 1) the entire 3.3 minute clip, 2) only the sound from the clip, 3) only subtitles and images from the clip, 4) only the prosodic intonations (with sounds garbled) and the images from the clip, 5) only subtitles, 6) only prosodic intonations, 7) only images, and 8) none of the above (controls who only watched the 8 minute

context clip). Compared with group 1 which received information through all three modalities and answered 87% of the questions correctly vs. group 8 that received no relevant information and scored 38%, all of the groups that had access to the verbal information (2, 3, and 5) fared best, answering in the 70–74% range. Participants that received only visual information (6) scored at about 62%, and those with only prosodic input scored 51%. Oddly, group 4, which had access to both prosody and images scored only 51% as well, probably due to the fact that it is a very unusual task to attempt to integrate images and prosodic intonations without the corresponding linguistic segments. Kibrik's conclusion is that the various channels of communication are not merely additive but also interact with each other and deserve much more study. At present his research group is collecting high-quality multimodal recordings of renarrations of Chafe's "The pear story" film. Early results (Kibrik and Fedorova 2016) show a clear but complex relationship between the purely visual stimulus of the film (which contains no speaking) and the segmentation of discourse into events. At the Multimodal Communication and Cognition Laboratory at Moscow State Linguistic University, headed by Alan Cienki (<http://scodis.com/?q=en/polimod-lab>), videos of narratives recorded in Russian are currently under analysis by a team of international linguists and gesture specialists. And recently the Distributed Little Red Hen Lab, which captures and archives newscasts for further study has begun collecting Russian, Polish and Czech (<http://babylon.library.ucla.edu/redhen/tweet?id=2016-09-05>), which means that there is now much more multimodal material available for future studies.

Bartoň et al. (2009: 166–68) have reported genre-related differences in the behavior of verbs (frequency of perfective vs. imperfective and in the most frequent lemmas for each aspect) in Czech, and it is likely that similar effects are relevant for other Slavic languages. Eckhoff et al. (submitted) have found that while the overall effect of aspect on the distribution of verb forms is quite robust, different genres are characterized by different verbs, and often the same verb will behave quite differently across genres. For example, the imperfective verb *provodit* 'carry out' patterns among ordinary, but not extreme imperfectives in Scientific-Technical prose but patterns with perfective verbs in both Journalistic prose and Fiction. However, the reason for this deviant patterning is also different across genres: in Fiction the indicative past is very prominent for this verb (accounting for 48.1% of its grammatical profile); however, it is the infinitive form of this verb that makes the strongest showing in Journalistic prose (36.3% of the profile).

#### 4. Variation across Time and Speakers

The last set of parameters for variation that we examine relate to the speakers themselves and the effects of time. Synchronically, speakers belong to dif-

ferent communities or different subgroups within their communities; here this will be reflected in a typological study and a gender study. Of course languages also change over time, both with respect to individual speakers, who acquire their languages as they mature, and with respect to entire communities, whose languages gradually change. This section will touch on areal typology, sociological effects on language (including gender), as well as language acquisition and diachronic change.

Arkad'ev (2015) undertakes an areal typology of verbal prefixation, examining a variety of factors across both Slavic and Caucasian languages. Using the NeighborNet clustering technique, Arkad'ev investigates both the relationships among languages and the relationships among factors. On the basis of these results, he posits two prototypes for the behavior of verbal prefixes (which he terms "preverbs"): a Slavic prototype in which it is possible to affix multiple prefixes, prefixes do not express deictic relationships, there is secondary imperfectivization, and perfective verbs cannot be combined with phasal verbs and cannot express durativity vs. a Caucasian prototype with the opposite characteristics. Within Slavic, Arkad'ev finds Slovene to be maximally similar to the other languages, forming together with Czech, Slovak, and Serbo-Croatian a core central European zone. This core zone has two near peripheries, one in the east consisting of Polish, Belarusian, Ukrainian and Russian, and one in the south encompassing Bulgarian and Macedonian. There are three possible causes for the areal phenomenon of perfectivizing preverbs: genetic relationships among languages, language contact, and universal typological tendencies. Among the Slavic languages, contact is a more important factor than in the Caucasus, and in both areas the semantic features of prefixes are more vulnerable to borrowing than other properties.

A study of gender by Kuznetsova (2015: Chapter 3) was facilitated by the fact that Russian marks gender on past tense verb forms. Using data from over six million past tense forms (belonging to 8340 lemmas) found in the modern subcorpus (after 1950) of the Russian National Corpus, Kuznetsova calculated the gender ratio (feminine:male) for verbs (excluding neuter forms and also verbs that do not have human subjects). She found that a typical Russian verb has approximately three times as many male forms as female forms, with nearly half of the verbs having between two and four times as many male forms. At the extremes of the distribution, Kuznetsova examined in detail the 100 most male and 100 most female verbs according to the gender ratio. Gender stereotypes are borne out by the grammatical profiles of verbs, but there are also some surprises. It is perhaps no surprise that the most male verbs are associated with professional occupations (such as *načal'stvovat'* 'be the boss'), physical strength (such as *kovat'* 'forge'), and negatively evaluated behavior (such as *p'janstvovat'* 'drink heavily'). While some of the most female verbs are associated as we might expect with motherhood (such as *zaberemet'* 'get pregnant') and household

tasks (such as *napeč'* 'bake'), Kuznetsova found in addition highly feminine verbs that refer to witchcraft (such as *nagadat'* 'tell fortunes') and to bird-like movements and sounds (such as *vyporxnut'* 'flit out' and *zaščebetat'* 'begin to twitter'). Thus from the perspective of Russian verbs, men are strong professionals who often behave badly, whereas women are birds who flit about their childrearing and housekeeping tasks, but may also have nefarious connections to the supernatural.

The acquisition of Russian aspect has been approached from the perspective of cognitive linguistics in a number of works by Stoll, beginning with her dissertation (Stoll 2001); however, here we focus on just one article about how children sort perfective verbs from imperfective verbs. Stoll and Gries (2009) looked at the distribution of perfective vs. imperfective verbs across past vs. non-past tense in the speech of child learners of Russian as compared to their adult caregivers. They used the Cramer's V statistic to measure the association strength between tense and aspect (manifested as a preference of perfective verbs for past tense but as a preference of imperfective verbs for non-past tense). Their finding was that whereas the speech of both children and adults shows this association, it is considerably more marked in the speech of children, who are initially very conservative and then gradually approach the norms of adults. In other words, the youngest children (ages 2–3) start out by using perfective verbs in the past tense and imperfective verbs in the non-past tense and then begin to loosen up this constraint.

Voejkova et al. (2015) also compare the language of children with that of adults, this time with respect to the use of adjectives. Their material is the spontaneous speech of children (age 1;5–3;0) with their families, which corresponds to the first year when children produce adjectives in Russian. Adjectives are of relatively low frequency and inherently rather abstract, since they implicitly require the comparison of a number of objects in order to arrive at a qualitative characteristic. It is thus no surprise that adjectives are typically learned later than nouns and verbs, and also unlike nouns and verbs, the understanding of adjectives does not precede their productive use by children. Voejkova and her team find that the input is adequate and plentiful, and furthermore that it is structured in a way that helps to establish the relevant cognitive categories. In that first year of adjective acquisition, the focus is on visible categories, such as: color (*belyj* 'white'), size (*malen'kij* 'small'), movement (*bystryj* 'fast'), distance (*dalekij* 'far'), and direction (*levyj* 'left'). Adjectives for touch and taste are less evident but tend to come in antonym pairs (*sladkij* 'sweet'/ *kislyj* 'sour', *suxoj* 'dry'/ *mokryj* 'wet'). This pattern mirrors the cognitive development of children, with visual discrimination coming before discrimination in other senses. The acquisition of adjectives is strongly supported by context, familiarity with relevant objects, and the use of qualitative questions (*Kakaja mašina?* 'What kind of car is it?'). This study supports the

idea that language acquisition is a direct product of the known developmental properties of the brain.

Since there is a whole article in this issue devoted to Slavic historical linguistics (Nesset this issue, 441–464), and a good deal of the work represented there takes the perspective of cognitive linguistics, we will approach diachrony here instead in terms of an ongoing change, namely the change in government from genitive to accusative for objects of verbs like *bojat'sja* 'be afraid' in Russian. Kuznetsova and Nesset (Kuznetsova and Nesset 2015, Nesset and Kuznetsova 2015) present both corpus data and an experiment that reveal the influence of various factors on this gradual language change, including: individuation of the object, grammatical voice, frequency, verb semantics, and register. Although the use of accusative is still relatively rare (under 10% for most verbs in the study), it was already evident in the mid 1800s and is still increasing, and this trend is strongest when the object is an animate noun or a proper noun (thus highly individuated) and in less restrictive registers (such as in newspapers). The construction grammar model enables the authors to represent the complex interplay of factors in this ongoing change.

## 5. Conclusion

In this discussion we have outlined what we consider to be some major contributions of cognitive linguistics to Slavic linguistics. These contributions stem in one way or another from the emphasis in cognitive linguistics on category structure (e.g., radial categories). This approach to the semantic meaning of grammatical and lexical categories has produced advances in our understanding of major Slavic grammatical categories, including case and aspect. Cognitive linguistics is particularly suited to the analysis of lexico-grammatical units such as prefixes, and it should come as no surprise that cognitive linguistics has taken a leading role in a revived study of prefixes over the last 30 years.

In general, cognitive linguistics is eminently suited to describe and explain variation in linguistic systems and communities: variation in the meanings of linguistic units as well as in the forms expressing a given meaning; variation across modalities and genres; and last but not least, variation across time and speakers. The compatibility of the approach of cognitive linguistics to meaning with statistical approaches in the description and explanation of variation has been demonstrated in recent years. We look forward to more advances in the years to come.

## References

- Antas, Jolanta. (2001) "Co mówią ręce. Wprowadzenie do komunikacji niewerbalnej". Renata Przybylska and Wiesław Przyczyna, eds. *Retoryka dziś*, 437–59. Kraków: Wydawnictwo Poligrafia salezjańska.
- Apresjan, Jurij D. (1995) "Obraz čeloveka po dannym jazyka: Popytka sistemnogo opisanija". *Voprosy jazykoznanija* 1: 37–67.
- Arkad'ev, Petr M. (2015) *Areal'naja tipologija prefiksals'nogo perfekтива*. Moscow: Jazyki slavjanskoj kul'tury.
- Bartoň, Tomáš, Václav Cvrček, František Čermák, Tomáš Jelínek, and Vladimír Petkevič. (2009) *Statistiky češtiny*. Prague: Nakladatelství Lidové Noviny.
- Będkowska-Kopczyk, Agnieszka. (2004) *Jezikovna podoba negationih čustev v slovenskem jeziku*. Ljubljana: Študentska založba.
- Christensen, Jason Heath. (2011) *The prefix PO- and aspect in Russian and Polish: A cognitive grammar account*. Ph.D. dissertation, University of Kansas.
- Clancy, Steven J. (2006) "The topology of Slavic case: Semantic maps and multidimensional scaling". *Glossos* 7: 1–28.
- Dąbrowska, Ewa. (1997) *Cognitive semantics and the Polish dative*. Berlin: Mouton de Gruyter.
- Dickey, Stephen M. (2000) *Parameters of Slavic aspect: A cognitive approach*. Stanford, CA: CSLI Publications.
- . (2005) "S-/Z- and the grammaticalization of Slavic aspect". *Slovene linguistic studies* 5: 3–55.
- . (2007) "A prototype account of the development of delimitative PO- in Russian". Dagmar Divjak and Agata Kočańska, eds. *Cognitive paths into the Slavic domain*. Berlin: Mouton de Gruyter, 326–371.
- Dickey, Stephen M. and Julie Hutcheson. (2003) Delimitative verbs in Russian, Czech and Slavic. Robert A. Maguire and Alan Timberlake, eds. *American contributions to the 13th International Congress of Slavists. Volume 1: Linguistics*. Bloomington: Slavica, 23–36.
- Dickey, Stephen M. and Laura A. Janda. (2009) "Xoxotnul, sxitiril: The relationship between semelfactives formed with -nu- and s- in Russian". *Russian linguistics* 33(3): 229–48.
- Dickey, Stephen M. and Laura A. Janda. (2015) Slavic aspectual prefixes and numeral classifiers: Two kinds of lexico-grammatical unitizers. *Lingua* 168: 57–84. DOI: 10.1016/j.lingua.2015.09.005
- Divjak, Dagmar. (2011) *Structuring the lexicon: A clustered model for near-synonymy*. Berlin: De Gruyter Mouton.
- Eckhoff, Hanne M. and Laura A. Janda. (2014) Grammatical profiles and aspect in Old Church Slavonic. *Transactions of the Philological Society* 112(2): 231–58.



- Eckhoff, Hanne M., Laura A. Janda, and Olga Lyashevskaya. (Submitted) "Predicting Russian aspect by frequency across genres". Unpublished ms., University of Tromsø
- Endresen, Anna. (2014) *Non-standard allomorphy in Russian prefixes: Corpus, experimental, and statistical exploration*. Ph.D. dissertation, University of Tromsø.
- Fidler, Masako Ueda. (2014) *Onomatopoeia in Czech: A conceptualization of sound and its connections to grammar and discourse*. Bloomington, IN: Slavica.
- Fortuin, Egbert. (2000) *Polysemy or monosemy: Interpretation of the imperative and the dative-infinitive construction in Russian*. Ph.D. dissertation, University of Amsterdam.
- Goldberg, Adele E. (2006) *Constructions at work: The nature of generalizations in language*. Oxford: Oxford University Press.
- Grigoréva, Svetlana A., Nikolaj V. Grigor'ev, and Grigorij E. Krejdlin. (2001) *Slovar' jazyka russkix žestov*. Moscow: Jazyki russkoj kul'tury.
- Jakobson, Roman (1959) "Boas' view of grammatical meaning". Walter Goldschmidt, ed. *The anthropology of Franz Boas: Essays on the centennial of his birth*, 139–45. Reprinted in Jakobson, Roman (1971) *Selected Writings*. Vol. 2, The Hague: Mouton, 489–96.
- Janda, Laura A. (1993) *A geography of case semantics: The Czech dative and the Russian instrumental*. Berlin: Mouton de Gruyter. [*Cognitive linguistics research*, 4.]
- . (2002) Cases in collision, cases in collusion: The semantic space of case in Czech and Russian. Laura A. Janda, Steven Franks, and Ronald Feldstein, eds. *Where one's tongue rules well: A Festschrift for Charles E. Townsend*. Columbus, OH: Slavica, 43–61.
- . (2004) A metaphor in search of a source domain: The categories of Slavic aspect. *Cognitive linguistics* 15(4): 471–527.
- Janda, Laura A. and Steven J. Clancy. (2002) *The case book for Russian*. Bloomington, IN: Slavica.
- . (2006) *The Case Book for Czech*. Bloomington, IN: Slavica.
- Janda, Laura A. and Valery Solovyev. (2009) What constructional profiles reveal about synonymy: A case study of Russian words for sadness and happiness. *Cognitive linguistics* 20(2): 367–93.
- Janda, Laura A., Anna Endresen, Julia Kuznetsova, Olga Lyashevskaya, Anastasia Makarova, Tore Nessel, and Svetlana Sokolova. (2013) *Why Russian aspectual prefixes aren't empty: Prefixes as verb classifiers*. Bloomington, IN: Slavica.
- Kemmer, Suzanne. (1993) *The middle voice*. Philadelphia: John Benjamins.
- Kibrik, Andrej. (2010) Mul'timodal'naja lingvistika. Yuri I. Alexandrov, Valery D. Solovyev, eds. *Kognitivnyye issledovanija–IV*. Moscow: IP RAN, 134–52.

- Kibrik, Andrej A. and Vera Podlesskaja, eds. (2009) *Rasskazy o snovidenijax: Korpusnoe issledovanie ustnogo ruskogo diskursa*. Moscow: Jazyki slavjanskix kul'tur.
- Kibrik, Andrej A. and O. V. Fedorova. (2016) "Vlijanie struktury zritel'nyx stimulov na sobytijnuju strukturu diskursa". Presentation at *Sobytija v komunikacii i kognicii*. Moscow. Available at: <http://www.scodis.com/files/documents/%D1%82%D0%B5%D0%B7%D0%B8%D1%81%D1%8B.pdf>
- Kuznetsova, Julia. (2015) *Linguistic profiles: Going from form to meaning via statistics*. Berlin: De Gruyter Mouton. [Cognitive linguistics research, 53.]
- Kuznetsova, Julia, and Tore Nessel. (2015) In which case are Russians afraid? *Bojat'sja* with genitive and accusative objects. *Journal of Slavic linguistics* 23(2): 255–83.
- Langacker, Ronald W. (2006) On the continuous debate about discreteness. *Cognitive linguistics* 17: 107–51.
- . (2013) *Essentials of cognitive grammar*. Oxford: Oxford University Press.
- Leinonen, Marja. (1982) *Russian aspect, "temporal'naja lokalizacija", and definiteness/Indefiniteness*. Ph.D. dissertation, University of Helsinki.
- Makarova, Anastasia. (2014) *Rethinking diminutives: A case study of Russian verbs*. Ph.D. dissertation, University of Tromsø.
- Maslov, Jurij S. (1984) *Očerki po aspektologii*. Leningrad: Leningrad State University.
- Mehlig, Hans Robert. (1996) Some analogies between the morphology of nouns and the morphology of aspect in Russian. *Folia linguistica* 30(1–2): 87–109.
- Maysak, Timur A. and Ekaterina V. Rakhilina, eds. (2007) *Glagoly dviženija v vode: leksičeskaja tipologija*. Moscow: Indrik.
- Moulton, Erin. (2015) *Rethinking reflexivity: Se- verbs in Russian and Bosnian/Croatian/Serbian*. Ph.D. dissertation, University of Kansas.
- Nessel, Tore. (2008) *Abstract phonology in a concrete model: Cognitive linguistics and the morphology-phonology interface*. Berlin: Mouton de Gruyter.
- Nessel, Tore and Julia Kuznetsova. (2015) Constructions and language change. From genitive to accusative objects in Russian. *Diachronica* 32(3): 365–96.
- Rakhilina, Ekaterina V. (2000) *Kognitivnyj analiz predmetnyx imen: Semantika i sočetaemost'*. Moscow: Russkie slovari.
- Say, Sergey. (2013) On the nature of dative arguments in Russian constructions with 'predicatives'. Irina Kor Chahine, ed. *Current studies in Slavic linguistics*. Amsterdam: Benjamins, 225–45.
- Šarić, Ljiljana. (2008) *Spatial concepts in Slavic: A cognitive linguistic study of prepositions and cases*. Wiesbaden: Harrassowitz Verlag.
- . (2014) *Prostor u jeziku i metafora: Kognitivnolingvističke studije o prefiksima i prijedlozima*. Zagreb: Naklada Jesenski i Turk.
- Shull, Sarah. (2003) *The experience of space: The privileged role of spatial prefixation in Czech and Russian*. Munich: Otto Sagner.

- Slobin, Dan I. (1996) From 'thought and language' to 'thinking for speaking'. John. J. Gumperz and Stephen C. Levinson, eds. *Rethinking linguistic relativity*. Cambridge: Cambridge University Press, 70–96.
- Šmelev, Aleksej D. (2002) *Russkaja jazykovaja model' mira*. Moscow: Jazyki slavjanskoj kul'tury.
- Sokolova, Svetlana, Olga Lyashevskaya, Laura A. Janda. (2012) The locative alternation and the Russian 'empty' prefixes: A case study of the verb *gruzit'* 'load'. Dagmar Divjak and Stefan Th Gries, eds. *Frequency effects in language representation*. Berlin: de Gruyter Mouton, 51–86. [*Trends in Linguistics. Studies and Monographs*, 244.2.]
- Stoll, Sabine. (2001) *The acquisition of Russian aspect*. Ph.D. dissertation, University of California, Berkeley.
- Stoll, Sabine and Stefan Th. Gries. (2009) How to measure development in corpora: An association strength approach. *Journal of child language* 36(5): 1075–90.
- Tabakowska, Elżbieta, ed. (2000) *Brygida Rudzka-Ostyn. Z rozważań nad kategorią przypadku*. Kraków: TAIWPN UNIVERSITAS.
- Talmy, Leonard. (2000) *Toward a cognitive semantics*. Vol. 2. *Typology and Process in Concept Structuring*. Cambridge, MA: MIT Press.
- Voejkova, Maria D., Viktorija V. Kazakovskaja, Dar'ja N. Satjukova. (2015) "Semantika prilagatel'nyx v reči vzroslyx i detej". A. A. Kibrik, A. D. Košelev, A. V. Kravčenko, Ju. B. Maurova, and O. V. Fedorova, eds. *Jazyk i mysl': Sovremennaja kognitivnaja lingvistika*. Moscow: Jazyki slavjanskoj kul'tury, 488–540.
- Zalazińska, Aneta. (2001) *Schematy myśli wyrażane w gestach: Gesty metaforyczne obrazujące wybrane abstrakcyjne relacje i zasoby podmiotu mówiącego*. Kraków: Universitas.
- Zaliznjak, Anna A., Irina B. Levontina, and Aleksej D. Šmelev. (2005) *Ključevye idei russkoj jazykovoj kartiny mira*. Moscow: Jazyki slavjanskoj kul'tury.

Laura A. Janda  
 UiT, The Arctic University of Tromsø  
 HSL fakultet, Universitetet i Tromsø,  
 N-9037, Tromsø, Norway  
 laura.janda@uit.no

Stephen M. Dickey  
 Department of Slavic Languages and Literatures  
 University of Kansas  
 1445 Jayhawk Blvd  
 Lawrence, KS 66045, USA  
 smd@ku.edu

