

In and Out of Context: Context, Redundancy, and Construal in Language

Background and status of knowledge

“I model jeans.” Or is it “genes”? Is this a statement about work in a laboratory or in the fashion industry? And how do I know that *model* is a verb here, rather than a noun, as in *fashion model*? Context matters. The *In and Out of Context* project examines the effects of contextual and non-contextual factors on the linguistic behavior of native speakers in order to create accurate theoretical models and effective learning tools. *In and Out of Context* focuses on a poorly understood yet obligatory distinction in Russian and employs corpus data analysis, experimentation, and machine learning to model that distinction. The methods and results can be implemented across the modeling and teaching of all languages.

Russian aspect is obligatory

In Russian, every time you use a verb, you have to choose between two forms: a perfective or an imperfective. Since most sentences have at least one verb, this is a pervasive phenomenon. This distinction (perfective vs. imperfective) is called **aspect** and its meaning and use defy accurate description. A vast scholarly literature¹ is devoted to Russian aspect, which is routinely listed as the greatest single obstacle facing second language learners of Russian.² Although aspect is found in many of the world’s languages, it is usually restricted to only some of a verb’s forms (as in Spanish, for example, where it appears only in the past tense). However, in Slavic languages like Russian, aspect is obligatory for all forms of verbs. One could say that perfective verbs describe situations as complete events, while imperfective verbs describe situations as ongoing or repeated processes, but this is a gross oversimplification.

There are both non-contextual and contextual cues for aspect

There are still many mysteries about how native speakers of Russian learn and use the category of aspect, and how this category can be taught to second language learners. A multitude of **cues** can indicate aspect in Russian, including an elaborate system of morphological cues that does not depend on context but unfortunately also has numerous exceptions). **Context** additionally delivers both **grammatical cues** (syntactic context of grammatical constructions) and **situational cues** (interpersonal communication known as deixis and pragmatics). The quantification of contextual cues and their interaction with non-contextual cues is precisely the mission of *In and Out of Context*.

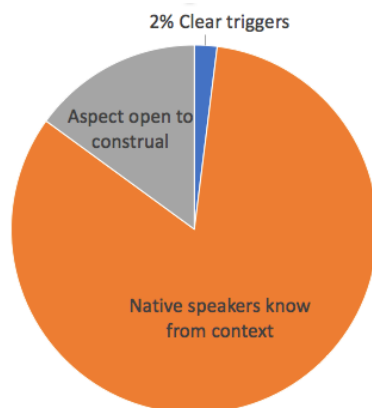


Figure 1: Context and use of aspect in

The 2% problem

Descriptive grammars of Russian list dozens of adverbs and other “triggers” that indicate aspect with fairly good reliability (around 96%). But these **triggers only work when they are available**. Our team has discovered (Reynolds 2016) that even when all known triggers are taken in aggregate, they are relatively rare in actual language use, appearing in association with **only about 2%** of verbs in corpus language samples (see Figure 1).³ This suggests that the known explicit contextual cues represent only the tip of the iceberg. This is a serious problem because textbooks and language courses present Russian aspect in terms of such triggers. This means that, as linguists and as instructors, we fail to represent 98% of the relationship of context to aspect.

How do native speakers use context?

Context is messy territory. By revealing the relationship of context to ambiguity, we can transform theoretical models of language and pedagogical materials. Context is formed by an intricate web of grammatical constructions, words with their meanings and frequencies, the speaker’s attitude toward the hearer and their feelings about the message, and even the physical setting where something is said. Context is also the cradle of meaning. It is from context that children discover concepts, since they don’t learn their native languages by reading definitions out of a dictionary. Or, as Firth (1957) famously put it: “You shall know a word by the company it keeps”. There is ample evidence that children are sensitive to statistical

¹ See references in Janda 2004 & 2007, Janda et al. 2013, and Eckhoff et al. 2017.

² Cf. Offord 2005, Andrews et al. 1997, Cubberly 2002, Martelle 2011.

³ A corpus is a collection of texts designed to represent authentic language use for the purpose of scientific investigation.

tendencies in the language they hear from their caregivers.⁴ From the perspective of corpus linguistics, context has been approached by means of behavioral profiles of words,⁵ examining a wide range of linguistic factors, among them morphological, syntactic, semantic and lexical factors. Linguistic profiling has been further honed in relation to Russian in the CLEAR (Cognitive Linguistics: Empirical Approaches to Russian) research group led by Janda and Nessel at UiT.⁶ Context is usually processed effortlessly by humans, who have no trouble using it to sort out ambiguities like *model* as a verb versus as a noun, as in our example above. But context is a major stumbling-block for computational models of language that struggle with rampant ambiguities of this type in human language. Using UDAR (cf. Hypothesis 1), we have estimated that 45% of words in running text in Russian are potentially ambiguous (see Figure 2), a problem that is solved by native speakers by reference to context. We are still only scratching the surface of how context works and how it interacts with other cues.



Figure 2: Ambiguity in Russian: выпей! (*vyпей!*) means both ‘drink up!’ and ‘bitterns (a type of bird)!’ (genitive/accusative plural)

Massive redundancy in language

It is known that language is a highly redundant phenomenon, and that this redundancy facilitates communication even in situations where the input is distorted.⁷ *In and Out of Context* will reveal how redundancy works and how native speakers use multiple signals from their inputs, even when some signals are missing or contradictory. In contexts where native speakers agree that there is only one correct aspect, one could argue that the non-contextual cues are redundant, since the aspect is clear without them. It might seem surprising that there are no indicators of Russian aspect that are 100% reliable. It is likely that native speakers combine cues to select aspect, but there is much we do not know about how redundancy works. *In and Out of Context* will address this enigma.

Construal: when the same situation can be expressed in different ways

We aim to discover how native speakers navigate construal so that we can guide learners to acquire a similar capacity to understand and use complex categories like aspect. In addition to contexts that require or strongly prefer one aspect, there are some contexts where native speakers readily accept both aspects (see the grey zone in Figure 1). In these contexts, aspect is employed by speakers to put a slightly different “spin” on what is otherwise the same content. For example, one could say *V 2025 my otmetim / budem otmečat’ vos’midesjatiletie okončanija Velikoj Otečestvennoj Vojny*. ‘In 2025 we **will celebrate** the eightieth anniversary of the end of WWII’, using either the perfective *otmetim* (emphasizing a complete event) or the imperfective *budem otmečat’* (emphasizing a process). While in a sense, using language is always a process of construal (Langacker 2013), the specifics of contexts where languages allow alternate construals is understudied. These are likewise the contexts that bedevil learners of Russian, since it seems that there are “no rules” and native speakers are “going on their gut feelings”.

The knowledge frontier for Russian aspect

Prior to *In and Out of Context*, most research on the relationship of Russian aspect to context has focused on the 2% of cases with triggers represented by the blue sliver in Figure 1, plus vague generalizations that are of little use in language modelling and teaching like “perfective is for bounded unique events”, “imperfective is for unbounded, ongoing situations” (see Janda 2004 for an overview showing that such analyses yield little more than synonymous labels for “perfective” and “imperfective”).

In our pilot experiment, 500 native Russian speakers rated the acceptability of 1346 perfective and imperfective forms in six texts of 1100-1700 words each. Participants were asked to rate the acceptability of both perfective and imperfective forms of the verb, but they were not told what the original form was. Acceptability was rated as 0 = “impossible”, 1 = “acceptable”, 2 = “excellent”. In Figure 3 the ratings are

⁴ See Goldberg 2006 and references therein.

⁵ Firth 1957, Geeraerts et al. 1999, Gries & Divjak 2009.

⁶ Janda & Solovyev 2009, Nessel et al. 2011, Janda & Lyashevskaya 2011 & 2013, Eckhoff et al. 2014a-b, Eckhoff & Janda 2014, Kuznetsova 2015, Janda 2016.

⁷ Dahl 2004, Dąbrowska 2016 and references therein; cf. Chiari 2007 for definition of linguistic redundancy in terms of information theory and scholarly overview.

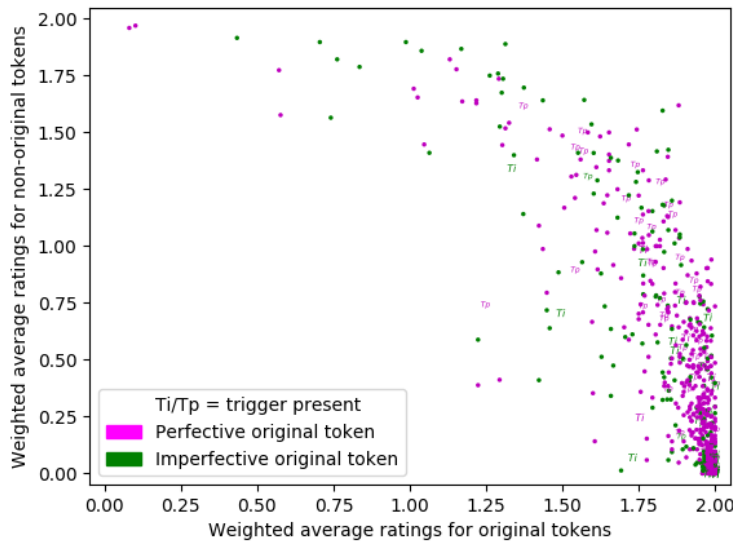


Figure 3: Comparison of weighted average ratings for original vs. non-original tokens

averaged over all participants, comparing the rating of the original token (the aspect that used in the original text) to the rating of the non-original token (the opposite aspect). The rating of the original token appears on the x-axis, with the rating of the alternative on the y-axis. Purple dots represent ratings of items where the original token was perfective, green represents original imperfectives. Ti and Tp replace dots in cases where there was a trigger for imperfective or perfective. Most of the data (81%) is located in the lower right quadrant, meaning that in the majority of cases respondents rated the original token highly and gave low ratings for the

non-original token. However, the majority of these items do not have triggers, and for most items the native speakers perform just as well regardless of whether there is a known trigger or not. These “triggerless” items in the lower right correspond to the orange sector in Figure 1 marked “Native speakers know from context”. The problem is, of course, that we don’t know how they do this. But at least now we can identify specific examples that illustrate this problem. Most of the rest of the data (17%) is in the upper right quadrant, where respondents gave both original and non-original tokens (both perfective and imperfective forms) acceptable-to-excellent ratings. The upper right quadrant of Figure 3 corresponds to the grey zone in Figure 1 marked “Aspect open to construal”. However, there are no distinct groups in this data: instead of distinct types, we find general statistical tendencies. A mixed-effects ordinal regression analysis of our data reveals that while many factors are significant, the factor visualized in Figure 3 (original vs. non-original aspect) is between 1 and 3 orders of magnitude larger in effect size than any other factor. Drilling deeper into our data, we also find that native speakers are more consistent in positive ratings for the original tokens, and less consistent in their ratings of the non-original tokens. But we still have a lot of explaining to do in order to account for all of our data. And even more if we want to account for how Russian aspect works.

Pushing the knowledge frontier for all languages

The techniques developed by *In and Out of Context* will be portable to other languages because all languages have a multitude of grammatical categories presenting choices that are dependent on context. Whereas grammars have traditionally claimed that there are strict criteria and rules governing such choices, since the advent of vast language data resources on the internet and sophisticated statistical software, linguists have increasingly become aware that language is more about statistical tendencies than about categorical rules. Bresnan (2016: 606-607) summarized this sea change in linguistics in her acceptance speech for a Lifetime Achievement Award from the Association for Computational Linguistics: “we had been relying all along on inconsistent binary grammaticality judgments that can be manipulated by changing the probabilities of the contexts, and we had vastly underestimated the human language capacity”. Our team has undertaken a number of investigations of such statistically-driven choices in Russian, involving choices of prefixes or suffixes (Dickey & Janda 2009, Nessel et al. 2010 & 2011, Baayen et al. 2013, Janda et al. 2013), the Russian version of the locative alternation (Sokolova et al. 2012), and near-synonyms (Janda & Solovyev 2009, Divjak 2010), and our work aligns with similar issues pursued by other researchers (cf. Say 2013). Thus we have spent many years training up for *In and Out of Context*, which focuses on the more pervasive and complex relationship between alternate aspectual forms.

Approach: Usage-based model of language and construction grammar

In and Out of Context frames Russian aspect and the larger questions of context, redundancy and construal from the perspectives of the usage-based model of language and construction grammar, which both belong to cognitive linguistics.⁸ The usage-based model makes the minimal assumption that language is a network of generalizations that emerge from input and human experience via general cognitive mechanisms.

⁸ Cf. Janda 2015, Langacker 2013 and references therein.

Construction grammar⁹ makes two fundamental claims: 1) the construction (a form-meaning pairing, at any level of complexity) is the relevant unit of linguistic analysis, and 2) a language is a structured aggregate network of constructions, also known as a “constructicon”.

Hypotheses and methods

Our hypotheses address how the range of contextual cues available to native speakers can be discovered and made available also to non-native learners of language, who could thus attain a higher level of proficiency.

Hypothesis 1: It is possible to extract hitherto unknown contextual cues from language data

We have access to huge quantities of naturalistic language data, much of it with linguistic annotation, such as the Russian National Corpus (600M words), RuTenTen (10B words), ruWAC (1B words), and SynTagRus (a gold-standard corpus). At UiT we have developed UDAR, the only open-source freely available full-scale finite state transducer model of Russian that accounts for morphophonemic stress, making it possible to render linguistic annotation for any Russian text in digital format. Linguistic annotation combined with computational models of language like UDAR make it possible to extract rich information about the context of each given use of a verb, such as: the grammatical categories it expresses, the part-of-speech categories of its collocates (the words around it) and the grammatical categories that they express, the frequencies of the verb form and its collocates, etc. Statistical analysis facilitates identification of factors that point toward one interpretation over another, in our case of perfective over imperfective.

In 2016, CLEAR joined forces with colleagues in Russia, Sweden, and five other countries to build the world’s first constructicons. A constructicon is an ideal source of cleaned language data on the interaction of context with aspect. Work on the Russian Constructicon (an open-source, free and public electronic resource) is actively ongoing, with nearly 700 constructions already cataloged. Many grammatical constructions invoke contexts in which one aspect is preferred over the other, a fact reflected in constructicon entries. Furthermore, a constructicon is not just an inventory, since there are also “families” of constructions, giving structure to the patterns of grammatical context in the language.

Hypothesis 2: Experiments can confirm the validity of cues for native speakers

The value of the discovered contextual factors can be validated via experiments, which make it possible to investigate how native speakers (L1, for “native language”) react to both the aspect in the original text and the alternative, a type of information not available in a corpus. Our substantial pilot study described above is to our knowledge the first experiment to address the use of aspect in rich context in a large-scale fashion. Our data confirm the pattern visualized in Figure 1, indicating complex interactions of contextual cues, and show that while L1 speakers react fairly consistently to stimuli that match the original aspect in a context, they display high uncertainty about stimuli that do not match the original aspect. Further experiments will determine which contextual cues are decisive for native speakers.

Hypothesis 3: Machine learning can ferret out patterns that otherwise evade detection

Corpus and experimental data provide empirical evidence of cues associated with aspect, but do not directly indicate what information speakers rely on. However, this data can be used in machine-learning models that find connections between the intuitions of our experiment participants and the contextual cues that might motivate their intuitions. Machine learning builds on corpus and experimental results, making it possible to create and test models of how contextual cues interact with Russian aspect. The goal is to identify all the textual features that would allow a machine-learning model to perform as well as a native speaker in predicting aspect. The patterns of second language learners can also be modeled, making it possible to pinpoint differences between native speakers and learners. These findings will be highly relevant to topics in second language acquisition theory, such as order of acquisition.

Hypothesis 4: Gains in knowledge can be directly implemented in language teaching

We can test how second language learners (L2) react to the patterns that have been identified through corpus analysis, experiments and machine learning. This will indicate what contexts and which verbs are most challenging for learning aspect. Our findings will lead to pedagogical innovations that enable students to learn to use verbal aspect at a native-like level.

⁹ Cf. Goldberg 1995 & 2006, Kay & Fillmore 1999, and Croft 2001.

The project plan, project management, organisation and cooperation

In and Out of Context is organized in four Work Packages. While all team members in the project can contribute to each Work Package, there are two assigned point persons in charge of moving each package forward. Janda will be the Director in charge of overseeing all Work Packages, with Reynolds as Assistant Director. Monthly meetings of the full team will be conducted via Skype, five intensive weekend Team Seminars will bring the team together in person. The Work Packages overlap strategically, and Workshops segue between Work Packages and bring in our Scientific Advisory Board to evaluate the overall flow of the project. *In and Out of Context* closes with a Conference summarizing findings on the relationship of context to use of aspect by native speakers (L1) and non-native learners (L2).

Year 1		Year 2		Year 3		Year 4	
WP1: Corpus and Constructicon Analysis Point Persons: Eckhoff & Nesset Assisted by: PostDocs & PhD student							
	WP2: L1 Experiments Point Persons: Janda &Lyashevskaya Assisted by: PostDocs & PhD student						
	Workshop 1	WP3: Machine Learning Point Persons: Divjak & Tyers Assisted by: PostDocs & PhD student					
				WP4: L2 Resources / Experiments Point Persons: Janda & Reynolds Assisted by: PostDocs & PhD student			
Team Seminar 1		Team Seminar 2	Team Seminar 3	Workshop 2	Team Seminar 4	Team Seminar 5	Conference L1 vs. L2 aspect

Figure 4: Gantt Chart for *In and Out of Context*

The *In and Out of Context* Work Packages address our four Hypotheses.

Work Package 1: Corpus and Constructicon

“Extract contextual cues from naturalistic and cleaned language data”

Analysis tools will automatically extract features of each word and sentence in corpus texts and Constructicon data. This results in fine-grained descriptions of each sentence, including word frequency (both alone and in collocation with other words), semantic roles and discourse functions, and also native speakers’ strategies in relation to genre. Based on these primary features, we can extract secondary features from corpus and Constructicon data, such as the probability that a particular word or construction will occur with imperfective or perfective verbs.

Work Package 2: Experiments with native speakers (L1)

“Reveal use of contextual cues by native speakers through experiments”

We will conduct a series of experiments using lexical decision, self-paced reading tasks, and acceptability ratings. Lexical decision and self-paced reading tasks measure how long it takes a participant to react to words, telling us how familiar or appropriate those words are both in and out of context (which can serve as a prime), and the context can also be adjusted to various “window” sizes (the number of words of context, also known as n-grams). For example, we can discover whether L1 speakers react more rapidly to verbs when they have received a cue to aspect (a prime), and whether they react more slowly when the inappropriate aspect appears. Acceptability scores measure L1 speakers’ conscious ratings of “how good” a given word sounds in a given context, and context can be manipulated in experiments. Collectively these experiments will make it possible for us to determine which cues to aspect are most decisive for L1 speakers.

Work Package 3: Machine learning

“Apply machine learning to find hidden patterns in context and determine which patterns matter”

We will use machine learning to refine and model our findings of contextual effects on aspect resulting from WP1 and 2. We intend to train models that output results comparable to human raters, exhibiting high confidence for examples where native speakers are unanimous, and lower confidence for examples where native speakers are divided. By manipulating the contextual information available to each model, we can identify which contextual information is sufficient for the task. Thus we can identify the contextual features that can help language learners perform at a native level. For example, if we can make a machine-learning model behave like an intermediate-level language learner by withholding a particular contextual feature, that

would lend support to the conclusion that awareness of that particular feature is what distinguishes intermediate language learners from native speakers on this task.

Work Package 4: Resources for and experiments with second-language learners (L2)

“Implement findings from WP1-3 in building and testing resources for language learners”

The results of corpus, experimental, and machine learning analysis will enhance the representation of aspect in the Constructicon, a free open-source resource for learners and instructors. In addition to the Russian Constructicon, a battery of interactive modules targeting the mastery of contextual patterns in relation to aspect will be added to UiT’s Russian “Oahpa!” electronic learning platform as well as the Strategic Mastery of Russian Tool (SMARTool, funded in a project beginning in 2018), resources that are freely available to the public. Our experiments will focus primarily on forced-choice tests, lexical decision tasks, and pre- and post-testing of learning modules, and in addition we can run experiments like those in WP2 to compare L1 vs. L2 reactions. Forced-choice tests will ask learners to decide whether perfective or imperfective aspect should be used in given contexts (manipulated in the experiment). These experiments need to be combined with lexical decision tasks to verify that L2 speakers can correctly identify verbs according to aspect.

Expertise

The *In and Out of Context* team represents vast and deep competence with regard to all facets of this project, as detailed in Table 1. This constellation of expertise ensures the success of *In and Out of Context*. All members of our project team have strong competence in disciplines relevant to this project, and all have published directly relevant articles in top peer-reviewed journals.

	Russian aspect	Grammatical constructions	Corpus data	Experimental methods	Machine learning	Statistical analysis	Pedagogical innovation
Laura A. Janda, PI	✓✓✓	✓✓	✓✓	✓		✓✓✓	✓✓✓
Dagmar Divjak (U Sheffield)	✓✓	✓✓	✓✓✓	✓✓	✓✓	✓✓✓	✓
Tore Nessel (UiT)	✓✓	✓✓	✓✓	✓✓		✓✓	✓✓✓
Olga Lyashevskaya (HSE Moscow)	✓✓	✓✓✓	✓✓✓	✓✓	✓	✓✓	✓
Hanne M. Eckhoff (Oxford U)	✓✓	✓✓	✓✓✓	✓	✓	✓✓	✓
Robert J. Reynolds (BYU)	✓	✓	✓✓	✓	✓✓	✓✓	✓✓
Francis M. Tyers (HSE Moscow)	✓	✓✓	✓✓	✓✓	✓✓	✓✓	✓

Table 1: Distribution of expertise across *In and Out of Context* team members

Key: ✓ = competence, ✓✓ = competence and publications in peer-reviewed venues, ✓✓✓ = recognized international leader

The expertise of the *In and Out of Context* team is enhanced by a Scientific Advisory Board with both national and international affiliations: Atle Grønn (U Oslo, Russian linguistics), Andrei Kutuzov (U Oslo, machine learning and language technology), Ekaterina Rakhilina (HSE Moscow, Russian linguistics and construction grammar), Alan Cienki (Vrije U Amsterdam & Moscow State Linguistic U; Russian linguistics, multimodal communication, language and cognition), Andrei Kibrik (Institute of Linguistics of the Russian Academy of Sciences, Russian and typological linguistics).

International cooperation

In and Out of Context constitutes a consortium of the institutions its team members are affiliated with: UiT (Janda & Nessel), Brigham Young U (Reynolds), Higher School of Economics in Moscow (Lyashevskaya & Tyers), U Sheffield (Divjak), and Oxford U (Eckhoff). *In and Out of Context* especially benefits from collaboration with U Sheffield’s *Out of Our Minds* project directed by Divjak and funded by the Leverhulme Trust to employ experimental and machine learning techniques to several linguistic categories, among them verbal aspect in Polish (which is partly similar to Russian). Experimentation will be facilitated by cooperation with Brigham Young U (BYU) and the National Research Higher School of Economics (HSE) in Moscow. BYU has one of the largest Russian programs in North America. HSE has recently launched an ambitious “Russian as a Foreign Language” program in addition to its numerous programs for native speakers, providing access to participants for experiments in WP2 & WP4.

Dissemination plan

Each Work Package will yield several scholarly articles, to be placed in top-ranked (and preferably open-access) journals, such as *Cognitive Linguistics*, *Studies in Language*, *Folia Linguistica*, *Functions of Language*, *Transactions of the Philological Society*, *Slavic and East European Journal*, *Voprosy jazykoznanija*, *Journal of Slavic Linguistics*, *Russian Linguistics*, *Annual Review of Cognitive Linguistics*, *Cognitive Science* (team members have previously published in all of these journals). Data and statistical

code associated with all publications will be publicly archived in TROLLing (the Tromsø Repository of Language and Linguistics, a professionally managed international resource based on Harvard's Dataverse Platform). Presentations will be made at conferences at the national and international level, such as NORKOG (the Norwegian Cognitive Linguistics Association), the Scandinavian Association for Language and Cognition, the International Cognitive Linguistics Conference, NODALIDA (Nordic Conference on Computational Linguistics), Dialogue (Computational Linguistics in Russia), COLING (International Conference on Computational Linguistics), and the Association for Computational Linguistics.

Communication with users

Promotional and instructional videos will be designed to engage learners and instructors in the Russian Constructicon, Russian "Oahpa!", and the SMARTool. These will be disseminated through professional networks of instructors and their channels of communication, such as the Fremmedspråksenteret of Norway and the American Association of Teachers of Slavic and East European Languages. In addition, social media sites will be established to broadcast results and updates, and to connect users to the project and to each other. These sites will encourage users to provide feedback and evaluation of our outputs, as well as recruit language learners for experiments on L2 acquisition of aspect.

References

- Andrews, Edna, Galina N. Aver'janova, Galina I. Pjadusova. 2001. *The Russian Verb: Form & Function*. Moscow: Russkij jazyk.
- Baayen, R. Harald, Anna Endresen, Laura A. Janda, Anastasia Makarova, and Tore Nessel. 2013. Making choices in Russian: Pros and cons of statistical methods for rival forms. *Russian Linguistics* 37:3, 253-291. DOI 10.1007/s11185-013-9118-6.
- Bresnan, Joan. 2016. Linguistics: The Garden and the Bush. *Computational Linguistics* 42(4):599-617.
- Chiari, Isabella. 2007. Redundancy Elimination: The Case of Artificial Languages. *Journal of Universal Language* 8, 7-38.
- Croft, W. 2001. *Radical construction grammar*. Oxford: Oxford University Press.
- Cubberly, Paul. 2002. Russian: A Linguistic Introduction. Cambridge: Cambridge U Press.
- Dąbrowska, Ewa. 2016. Cognitive Linguistics' seven deadly sins. *Cognitive Linguistics* 27:4, 479-491.
- Dahl, Östen. 2004. *The Growth and Maintenance of Linguistic Complexity*. Amsterdam: John Benjamins.
- Dickey, Stephen M. and Laura A. Janda. 2009. *Xoxotmul, sxitritl*: The relationship between semelfactives formed with -nu- and s- in Russian. *Russian Linguistics*, 33: 3, 229-248.
- Divjak, Dagmar. 2010. *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* Vol 112, Issue 2, pages 231-258.
- Eckhoff, Hanne M., Laura A. Janda, and Tore Nessel. 2014a. Old Church Slavonic byti Part One: Grammatical Profiling Analysis. *Slavic and East European Journal* 58.3, 482-497.
- Eckhoff, Hanne M., Laura A. Janda, and Tore Nessel. 2014b. Old Church Slavonic byti Part Two: Constructional Profiling Analysis. *Slavic and East European Journal* 58.3, 498-525.
- Eckhoff, Hanne M., Laura A. Janda, and Olga Lyashevskaya. 2017. Predicting Russian Aspect by Frequency Across Genres. *Slavic and East European Journal* 64:1 (2017), 844-875.
- Firth, John R. 1957. A synopsis of linguistic theory 1930-1955. *Studies in Linguistic Analysis*. Ed. John R. Firth. Oxford: Blackwell, 1-32.
- Geeraerts, D., S. Grondelaers, and D. Speelman. 1999. *Convergentie en divergentie in de Nederlandse woordenschat. Een onderzoek naar kleding- en voetbaltermen*. Amsterdam: Meertens Instituut.
- Goldberg, Adele. 1995. *Constructions: A construction grammar approach to argument structure*. Chicago: Chicago University Press.
- Goldberg, Adele E. 2006. *Constructions at Work: The Nature of Generalizations in Language*. Oxford: Oxford U Press.
- Gries, Stefan Th. and Dagmar Divjak. 2009. Behavioral profiles: a corpus-based approach towards cognitive semantic analysis. In: *New Directions in Cognitive Linguistics*. Eds. Vyvyan Evans and Stephanie S. Pourcel, Amsterdam: John Benjamins, 57-75.
- Janda Laura A. 2004. A metaphor in search of a source domain: the categories of Slavic aspect. *Cognitive Linguistics* 15:4, 471-527.
- Janda, Laura A. 2007. Aspectual clusters of Russian verbs. *Studies in Language* 31:3, 607-648.
- Janda, Laura A. 2013. Quantitative Methods in *Cognitive Linguistics*. In Laura A. Janda, ed. *Cognitive Linguistics: The Quantitative Turn. The Essential Reader*, 1-32. Berlin: De Gruyter Mouton.
- Janda, Laura A. 2015. Cognitive Linguistics in the Year 2015. *Cognitive Semantics* 1, 131-154.
- Janda, Laura A. 2016. "Linguistic profiles: A quantitative approach to theoretical questions". *Jazyk i metod* 3, 127-146.
- 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, 367-393.
- Janda, Laura A. and Olga Lyashevskaya. 2011. Grammatical profiles and the interaction of the lexicon with aspect, tense and mood in Russian. *Cognitive Linguistics* 22:4, 719-763.
- Janda, Laura A. and Olga Lyashevskaya. 2013. Semantic Profiles of Five Russian Prefixes: po-, s-, za-, na-, pro-. *Journal of Slavic Linguistics* 21:2, 211-258.
- Janda, Laura A., Anna Endresen, Julia Kuznetsova, Olga Lyashevskaya, Anastasia Makarova, Tore Nessel, Svetlana Sokolova. 2013. *Why Russian aspectual prefixes aren't empty: prefixes as verb classifiers*. Bloomington, IN: Slavica Publishers.
- Kay, Paul and Charles Fillmore. 1999. Grammatical constructions and linguistic generalizations: the what's X doing Y? construction. *Language* 75,1-34.
- Kuznetsova, Julia. 2015. *Linguistic profiles: Going from form to meaning via statistics (= Cognitive Linguistics Research 53)*. Berlin: De Gruyter Mouton.
- Langacker, Ronald W. 2013. *Essentials of Cognitive Grammar*. Oxford: Oxford U Press, 2013.
- Martelle, Wendy. 2011. *Testing the Aspect Hypothesis in L2 Russian*. Doctoral Dissertation, University of Pittsburgh.
- Nessel, Tore, Laura A. Janda, R. Harald Baayen. 2010. Capturing Correlational Structure in Russian Paradigms: a Case Study in Logistic Mixed-Effects Modeling. *Corpus Linguistics and Linguistic Theory* 6 (2010), 29-48.
- Nessel, Tore, Anna Endresen and Laura A. Janda. 2011. Two ways to get out: Radial Category Profiling and the Russian Prefixes vy- and iz-. *Zeitschrift für Slavistik* 56:4 (2011), 377-402.
- Offord, Derek. 2005. *Using Russian: A guide to contemporary usage*. Cambridge: Cambridge University Press.
- Reynolds, Robert J. 2016. *Russian natural language processing for computer-assisted language learning*. Doctoral Dissertation, UiT The Arctic University of Norway.
- Say, Sergey. 2013. On the Nature of Dative Arguments in Russian Constructions with 'Predicatives'. In Irina Kor Chahine (ed.), *Current Studies in Slavic Linguistics*, 225-245. Amsterdam: Benjamins.
- Sokolova, Svetlana, Olga Lyashevskaya, Laura A. Janda. 2012. The Locative Alternation and the Russian 'empty' prefixes: A case study of the verb *gruzit'* 'load'. In: D. Divjak & St. Th. Gries (eds.), *Frequency effects in language representation* (Trends in Linguistics. Studies and Monographs. 244.2), 51-86. Berlin: de Gruyter Mouton.