V zeleni drželi zeleni breg

Studies in Honor of Marc L. Greenberg

Edited by

Stephen M. Dickey &
Mark Richard Lauersdorf

Bloomington, Indiana, 2018

SLAVICA

ISBN 978-0-89357-491-8

Technical Editors: Phillip Weirich and Renata Uzzell

Cover Design by Tracey Theriault

Cover Image: Aleš Maučec (www.facebook.com/ales.maucec). Slovenia (Prekmurje, Goričko), Vučja Gomila—a small village (pop. 269) in the northeastern part of Slovenia near the border with Hungary.
© Aleš Maučec, permission granted by the author.

Library of Congress Cataloging-in-Publication Data

Names: Dickey, Stephen M., editor. | Lauersdorf, Mark Richard, editor. | Greenberg, Marc L., honoree.
Title: V zeleni drželi zeleni breg : studies in honor of Marc L. Greenberg / edited by Stephen M. Dickey, Mark Richard Lauersdorf.
Description: Bloomington, Indiana : Slavica Publishers, 2018.
Identifiers: LCCN 2018037963 | ISBN 9780893574918
Subjects: LCSH: Slavic languages. | Slavic languages, Southern.
Classification: LCC PG14.G54 V33 2018 | DDC 491.8/1--dc23
LC record available at https://lccn.loc.gov/2018037963

Slavica Publishers
Indiana University
1430 N. Willis Drive
Bloomington, IN 47404-2146
USA

[Tel.] 1-812-856-4186
[Toll-free] 1-877-SLAVICA
[Fax] 1-812-856-4187
[Email] slavica@indiana.edu
[www] http://www.slavica.com/

Contents

A Stranger in the Lexicon: The Aspectual Status of Russian *смочь* 'be able, manage (to)'

Laura A. Janda

Abstract: It has been claimed that Russian *смочь* 'be able, manage (to)' has a number of unusual properties relating to its expression of aspect and tense. A number of new kinds of data are brought to bear in this debate. This article compares *смочь* with its purported aspectual partner verb *мочь* using overall and longitudinal corpus data. It also compares the distribution of forms of *смочь* with those of other Russian verbs and reports on an experiment in which native speakers of Russian rated the acceptability of past tense *смог* in contexts where *мог* is attested. In addition, parallel corpus data is used to compare forms of Russian *смочь* with their translation equivalents in both Czech and Spanish. Collectively this data shows that *смочь* is arguably the most deviant purportedly perfective verb in Russian, and that it has shown a dramatic increase in frequency over the past century. However, it is not easy to identify the cause of this increase, nor to find strong support for the hypothesis that this is due to the expansion of nonpast forms of *смочь* to contexts where it merely expresses futurity.

1. Introduction

This article presents new evidence concerning the status of the Russian verb *смочь* 'be able, manage (to)', which seems to be an aspectual anomaly. This verb attracted some attention in the scholarly literature 15–20 years ago, but today there is more data available that can be brought to bear on this question.

Modality in Russian is peculiar because it is usually expressed by constructions involving adjectives and adverbs (like *должен* 'should', *надо* 'must', *нужно* 'necessary', *возможно* 'possible') or impersonal verbal constructions (like *придется* 'will have to', *понадобится* 'will be necessary', *надлежит* 'has to', *не спится* 'can't sleep'). Aside from *должен*, these means of expressing modality identify an experiencer marked in the Dative case, which may indicate a bias toward description of modal forces as externally imposed (but note that such a bias is difficult to prove, cf. Janda and Divjak 2008). In effect, Russian has only one modal verb that can take a nominative subject: *мочь* 'be able' (Divjak 2010: 76). This sets Russian apart from other Slavic languages which typically have, in addition to verbs derived from **moktī* 'be able', other modal

Stephen M. Dickey and Mark Richard Lauersdorf, eds. *V zeleni drželi zeleni breg: Studies in Honor of Marc L. Greenberg*. Bloomington, IN: Slavica Publishers, 2018, 105–26.

verbs related to **iměti* 'have', **morati* 'must', and/or a loan verb cognate with modern German *müssen* 'must'.

Furthermore, Russian *мочь* appears to have a perfective partner verb, *смочь*, with even more peculiar properties. *Смочь* was the topic of a series of works by Choi (1994, 1999) and Barentsen (2002), who advanced numerous claims about the behavior and status of this verb. Choi (1994) makes the following statements about *смочь*:

- non-past forms are used "to express future time of the situation of possibility, rather than to express its 'perfectivity'" (p. 169);
- past forms are "used to express the discourse function of sequentiality, rather than again to express the 'perfectivity' of the state of affairs" (p. 169);
- past forms are "combined exclusively with the perfective infinitive" (pp. 170–71), although combination is possible if the imperfective is an aspectually unpaired verb (imperfectivum tantum) or the conditional *бы* is present (p. 225);
- non-past forms are "used to express what the would-be future tense form of [*мочь*], which does not exist in Russian ([**буду мочь*]), is supposed to convey" (p. 171, cf. 217);
- while *мочь* doesn't usually occur in the infinitive, *смочь* can occur in the infinitive (with some restrictions) (p. 175).

Choi casts doubt on the assumption of "most dictionaries" (1994: 220) that *смочь* is the aspectual partner verb of *мочь*. He reasons that the use of non-past forms to express future due to the exclusion of **буду мочь* from Russian grammar is motivated by the fact that modals are inherently state verbs and therefore it is "inconceivable" that *смочь* could express the boundedness or totality that would be required of a perfective verb. Subsequently Choi (1999) amends this conclusion by claiming that *смочь* is a procedural semelfactive perfective in which the prefix *с-* has the same function as in verbs like *сглупить* 'do one foolish thing'.

Barentsen (2002), writing in reaction to Choi, presents different findings. Barentsen does not find it "inconceivable" that a modal verb might express perfective aspect since this is found in other languages, for example French *pouvoir* 'be able' appears in both perfective (*je pus*) and imperfective (*je pouvais*) past tense forms. Barentsen provides a couple of corpus examples that disprove Choi's claim that past forms of *смочь* occur only with perfective infinitives because in Barentsen's examples the verbs in question are neither aspectually unpaired nor collocated with *бы*. And Barentsen reports some further peculiarities of *смочь*:

- the frequency *смочь* of has grown remarkably (a ten-fold increase) in the past two centuries, and most of this gain comes from the use of non-past forms;
- the only potential cognates in Slavic are Ukrainian *змогти*, Bulgarian *смогна*, Macedonian *смогне*, and Czech *ztmoci*,[1] but it is not clear to what extent these correlate to *смочь*;
- translation equivalents of *смочь* in other Slavic languages tend to use imperfective equivalents of *мочь*;
- while *мочь* has no future forms in Russian, it does have future forms in Polish, Czech, Serbian, and Croatian.

Barentsen's conclusion is that *смочь* does indeed occupy a special position, but these are not just random facts. Instead these facts collectively point to systematic peculiarities of the meaning of Russian aspect, which is more categorical than in other Slavic languages.

While both Choi and Barentsen illustrate their claims with authentic examples, and, in the case of Barentsen, also with some corpus statistics, both the quantity of data and the means to analyze it have advanced dramatically in the intervening years. In particular, both the Russian National Corpus (RNC, ruscorpora.ru, last accessed November 2017) and the ParaSol corpus (Parallel Corpus of Slavic and Other Languages, Waldenfels 2011, parasolcorpus.org, last accessed November 2017) have become available in the meantime. It is therefore worth revisiting the behavior of *смочь* in light of these new data sources. In Sections 2-4 I will test and extend claims made by Choi and Barentsen by means of modern corpus data and also some experimental data, beginning first with data pertaining just to Russian (Section 2), followed by comparisons with the Slavic language that likely differs most from Russian in terms of aspect, namely Czech (cf. Dickey 2000, Section 3), and a comparison with a non-Slavic language that expresses aspect, namely Spanish (Section 4). I will not, however, engage in a detailed semantic analysis of individual examples.

2. Language-Internal Evidence: Russian Corpus and Experimental Data

I present three types of evidence documenting the behavior of *смочь* from the internal perspective of Russian. The first two types of evidence are based on data found in the Russian National Corpus (RNC), both of which examine the behavior of *смочь* in comparison with other Russian verbs. In these two studies, the measure of behavior is the grammatical profile, which is the relative frequency distribution of the inflected forms of a lexeme. In other words, we

[1] Note that this overview neglects cognates in other languages such as BCS, Slovene, and Polish.

look at how often the verb *смочь* appears in all of its forms (*смогу*, *сможешь*, etc.) and compare that frequency distribution with the frequency distribution of other verbs. In section 2.1 this comparison is made specifically with *мочь*, and additional RNC data is cited in relation to specific claims that have been made about *смочь*. Section 2.2 reports on a study done on the grammatical profiles of hundreds of high-frequency verbs across three genres, in which *смочь* was consistently found to behave in an aspectually anomalous fashion. Experimental data is presented in section 2.3, where we see how native speakers of Russian react to the use of *смог* vs. *мог* in the context of a narration.

2.1. The Grammatical Profile of *смочь* (Compared with *мочь*)

What can grammatical profiles and longitudinal statistics tell us about the relative distributions of inflected forms of *мочь* and *смочь* in both modern Russian and its recent history? Does this data corroborate claims made by Choi and Barentsen?

Table 1 presents data on the distribution of examples of *мочь* and *смочь* in the Russian National Corpus. This table shows both the raw numbers of attestations for each form ("# of examples"), as well as the percentage that each form represents in relation to the whole verb. The latter distribution of percentages is the grammatical profile of the verb. For both verbs, indicative forms predominate. For *мочь*, 99.45% of all forms are indicative (69.38% non-past, 30.08% past), while for *смочь*, 97.74% of forms are indicative (53.12% non-past, 44.62% past). Imperatives are quite rare for both verbs, as are gerunds and participles, aside from the present active participle *могущий*.

Recall that Choi claimed that *мочь* is extremely rare in the infinitive form, whereas *смочь* is less rare. The RNC data, on the contrary, shows no appreciable difference in the frequency of infinitives for these two verbs.[2] Choi (1994: 175) mentions that there are restrictions on *смочь* as an infinitive form, but offers only one concrete type, the *чтобы* clause. However, while many examples of *смочь* do occur in *чтобы* clauses, there are also many that don't, as in:

(1) Use of infinitive form ***смочь*** in *одно дело* + infinitive construction:

*Л*егко сказать, одно *д*ело использовать *л*ичное местоимения, а *д*ругое *д*ело, *д*ействительно, **смочь** мыслить от собственного «я» (*если* ты раньше этого не *делал*). [В. А. Подорога. Проект и опыт (2004)]

[2] A chi-squared test comparing the number of infinitives to the total number of forms for each verb yields the following result: X-squared = 5.5767, df = 1, *p*-value = 0.0182, Cramer's V = 0.003. In other words, the effect size (Cramer's V) falls two orders of magnitude below that of a reportable difference.

Table 1. Forms of *мочь* and *смочь* attested in the Russian National Corpus

	Form	# of examples	Percent	Form	# of examples	Percent
infinitive	*мочь*	537	0.06%	*смочь*	40	0.09%
1sg nonpast	*могу*	81 785	9.34%	*смогу*	4 282	10.00%
2sg nonpast	*можешь*	15 386	1.76%	*сможешь*	1 400	3.27%
3sg nonpast	*может*	383 082	43.73%	*сможет*	8 900	20.79%
1pl nonpast	*можем*	23 778	2.71%	*сможем*	2 305	5.38%
2pl nonpast	*можете*	20 138	2.30%	*сможете*	1 845	4.31%
3pl nonpast	*могут*	83 653	9.55%	*смогут*	4 492	10.49%
masc past	*мог*	130 552	14.90%	*смог*	10 082	23.55%
fem past	*могла*	46 563	5.31%	*смогла*	3 685	8.61%
neut past	*могло*	28 842	3.29%	*смогло*	392	0.92%
pl past	*могли*	57 535	6.57%	*смогли*	5 348	12.49%
2sg imper	*моги*	251	0.03%	*смоги*[3]	7	0.02%
2pl imper	*могите*	11	<0.01%	*смогите*	1	<0.01%
present active participle	*могущий*	3 576	0.41%	[NA]		
past active participle	*могший*	210	0.02%	*смогший*	32	0.07%
gerund	*могши*	200	0.02%	*смогши*	8	0.02%
	Total	**876 099**	**100%**	**Total**	**43 719**	**100%**

(1) 'It's easy to say that it's one thing to use a personal pronoun and another thing **to be** really **able** to think about one's own "I" (if you haven't done this before).'[4]

(2) Use of infinitive form ***смочь*** in infinitive-*то* + 1pl construction with reduplicated verb:

Смочь-то сможем, но это будет не слишком большой компенсацией за поражение умных. [Юлий Андреев, Валерий

[3] There were actually 9 attestations, but one was for the plural of *смог* 'smog' and another was a part of another word written out with hyphens in a song: *По-смоги-ить, кто в бога виру-е…* from Максим Горький. *Песня о слепых* (1901)

[4] Examples are from the RNC and ParaSol corpus, cited with their passports. All translations are mine.

> *Лебедев.* Моральный ум? (2003) // Интернет-альманах «*Лебедь*», 2003.10.19]
>
> 'Well, we can do it, but there won't be much compensation for defeating the intellectuals.'

Therefore I do not find support for Choi's claims concerning the distribution of the infinitives of *мочь* and *смочь*.

Recall also Choi's claim that past tense forms of *смочь* can be followed only by perfective infinitives, except in cases where the verb is an imperfectivum tantum or is collocated with conditional *бы*. To contest this claim, Barentsen (2002: 9) provides two corpus examples of *смог* followed by imperfective infinitives of aspectually paired verbs that are not collocated with *бы*. Today's RNC gives more support to Barentsen's argument. There are 701 examples of past tense forms of *смочь* immediately followed by an imperfective infinitive in the RNC -- hardly a rare occurrence as claimed by Choi. This is as opposed to 9,803 occurrences of past tense forms of *смочь* immediately followed by a perfective infinitive. In other words, approximately 7% of sequences with past tense forms of *смочь* followed by an infinitive involve an imperfective infinitive, and it is easy to find examples that do not follow Choi's stipulated restrictions, such as in:

(3) Use of ***смочь*** + imperfective infinitive of an aspectually paired verb and without *бы*:

> Однако, к счастью, сильного задымления там не наблюдалась [sic], и самолёты **смогли** взлетать и садиться строго по графику. [*Инна Левит.* Пожар на Шимякинском полигоне (2002) // «Вечерняя Москва», 2002.04.11]
>
> 'However, fortunately, no heavy smoke was observed and the airplanes **could** take off and land precisely according to schedule.'

Now recall Barentsen's (2002: 26-27) claim that there has been a dramatic increase in the use of *смочь* in all its forms over the past two centuries. This claim is based on a rather small sample of 5,000 pages of text for each half century and a total of 411 forms of *смочь* spread across the four time periods. The graphing functions available on the RNC page allow us to test this claim on the basis of much more data over the same time period, as shown in Figures 1 and 2. Both figures measure the frequency of forms of *смочь* per million words (the scale of the *y*-axis).

While Figures 1 and 2 confirm the overall gist of Barentsen's claim, they also give us much more detail. We see that the rise in frequency comes only in the 20th century. Furthermore, contra Barentsen, the difference in frequency growth is not related to the distinction of non-past vs. past, but rather to spe-

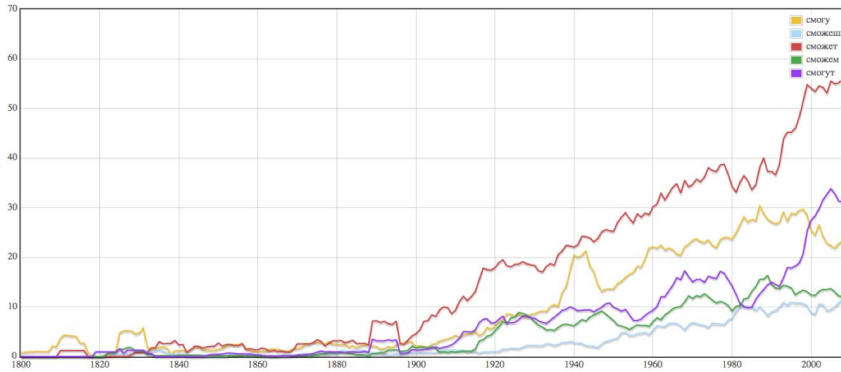

Figure 1. Frequency of non-past forms of *смочь* per million words 1800–2010

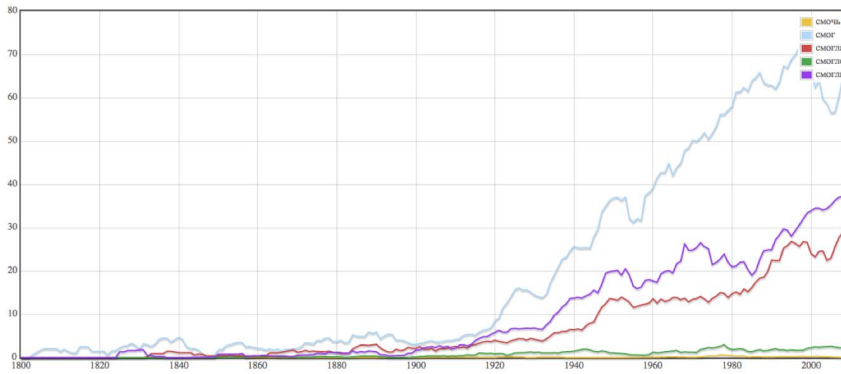

Figure 2. Frequency of past and infinitive forms of *смочь* per million words 1800–2010

cific forms: *сможет* and *смог* have shown the strongest growth, followed by (in decreasing order) *смогли*, *смогут*, *смогла*, and *смогу*. The remaining non-past forms have also made robust gains, but *смочь* and *смогло* have remained rather infrequent.

This trend begs the question of how it was motivated. Could it have something to do with changes in what *смочь* expresses, particularly in the forms *сможет* and *смог*? Choi states that *мочь* lacks a periphrastic future, a fact which is confirmed by modern corpus data. Could it be that *сможет* has moved in to take over uses previously expressed by a periphrastic future of *мочь*? This does not seem to be the case. Padučeva (2001) states that there

was no use of forms like **буду мочь* at Pushkin's time (early 19th century) either, and the RNC lists only four rather marginal examples, all from a very narrow time period (1894–98).[5] In other words, there is no substantial use of a periphrastic future in the 19th century that could have been taken over by nonpast forms of *смочь* in the 20th century. However, Padučeva (2001) also observes that the use of past tense forms of *смочь* was very rare in the early 19th century, and in contexts where today we use *смог*, the form *мог* appeared instead. Padučeva's observation is supported by Figure 3, where we see that the frequency of *мог* (and other past tense forms) has indeed dropped over the same time period. Thus we find some hints about the rise in use of past tense of *смочь*, where it seems to be taking over some of the uses of *мочь*, but no corresponding explanations for expansion of nonpast tense forms. Nonpast forms of *мочь* (not pictured) have, by contrast, remained rather steady in their frequency over time. I have more to say about the behavior of *смог* in Sections 2.3, 3, and 4.

[5] These four examples are:

(i) Тревожные мысли, что я не **буду мочь**, что мне преградят дорогу, что не дадут возможности принести ему малейшую пользу, бросали меня в глубокое уныние. [В. П. Авенариус. Чем был для Гоголя Пушкин (1895)]

'Alarming thoughts that I **will** not **be able**, that they will block me, won't give me a chance to assist him at all, threw me into a deep depression.'

(ii) Это уже столь весьма обольстительно сделалось в фантазии маменьки, что оне даже заплакали от счастия видеть меня в облачении в парчовом стихаре, наверно воображая меня уже малым чем умаленного от ангел и в приближении к наивысшему небу, откуда уже **буду мочь** кое-что и сродственникам своим скопнуть наземлю. [Н. С. Лесков. Заячий ремиз (1894)]

'It had already become so tempting in mother's mind that they even wept for joy when they saw me dressed in a brocade vestment, probably imagining me as a lesser angel approaching the highest heavens, from whence I **will be able** to kick something down to my relatives on earth. '

(iii) Скоро ли ты продерешься сквозь векселя, отчеты, разделы и т. п. и **будешь мочь** думать без выкладки на счетах и писать так, чтобы не мерещился двуглавый орел в заглавии листа? [Б. Н. Чичерин. Воспоминания (1894)]

'Will you soon get through all the bills, reports, clauses, etc. and you **will be able** to think without making calculations and write without a two-headed eagle looming at the head of the page?'

(iv) Авось не зажилит, а поплатится, когда **будет мочь**. [С. Т. Семенов. Алексей заводчик (1898)]

'Maybe she won't just take it, but will pay when she **will be able to**.'

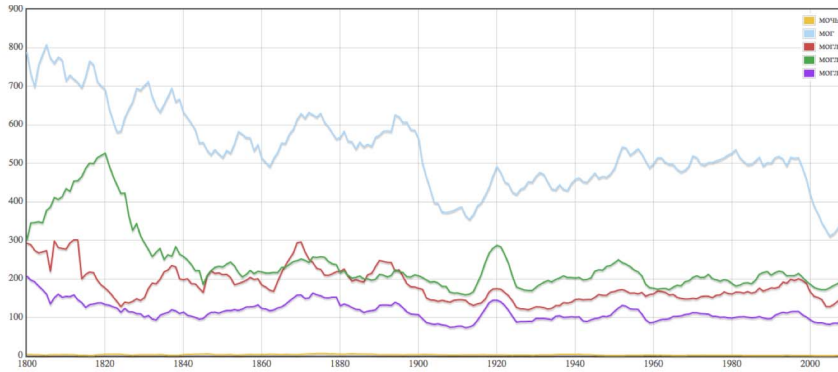

Figure 3. Frequency of past and infinitive forms of *мочь* per million words 1800–2010

2.2. The Grammatical Profile of *смочь* Compared with Other Verbs

How does the grammatical profile of *смочь* compare with other verbs, particularly in relation to verbal aspect? Can such data corroborate Choi's (1994) claim that *смочь* does not really mark perfective aspect, particularly in its nonpast forms?

Janda and Lyashevskaya (2011) documented a stark difference in the grammatical profiles of perfective vs. imperfective verbs based on a sample of 6 million verb forms from the Modern subcorpus of the RNC (1950–2007), visualized in Figure 4 on the following page. The grammatical profile of imperfective verbs is dominated by indicative nonpast forms (comprising 47.43% of their profile), while the grammatical profile of perfective verbs is dominated by indicative past forms (comprising 62.67% of their profile). Janda and Lyashevskaya (2011) showed the grammatical profiles in aggregate, which smoothed over individual differences between verbs. In other words, this study showed that it is possible to distinguish a group of perfective verbs from a group of imperfective verbs based on their grammatical profiles. However, it remained to be seen whether the grammatical profiles of individual verbs could be used to predict their aspect.

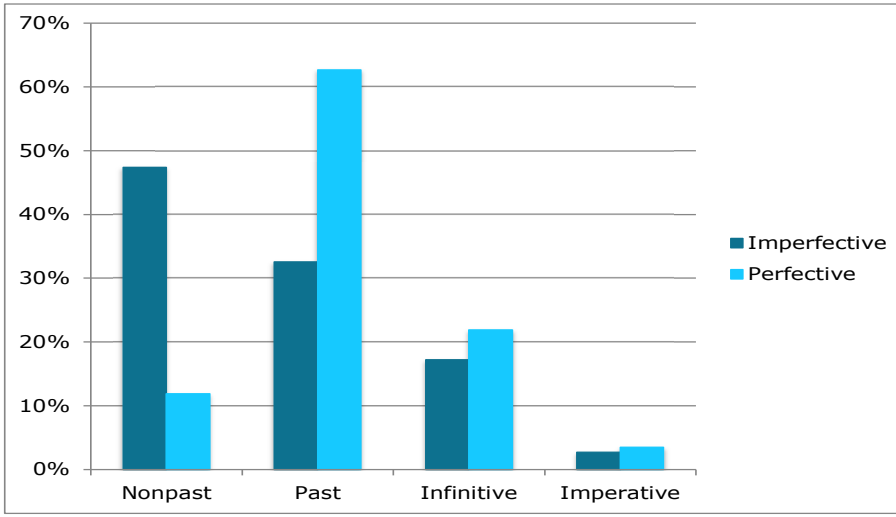

Figure 4. Aggregate grammatical profiles of 6 million imperfective vs. perfective verb tokens from the Modern subcorpus of the RNC, based on data in Janda and Lyashevskaya 2011.

Eckhoff et al. (2017) addressed the question of whether grammatical profiles can predict the aspect of individual verbs. They report on a study of the grammatical profiles of high-frequency (> 50 attestations) verbs from the manually disambiguated Morphological Standard of the Russian National Corpus (approximately 6 million words) representing the years 1991-2012. This study is stratified across three genres (journalism, fiction, scientific and technical writing) with same-sized samples (0.4 million words) for each. There were 185 verbs that crossed the frequency threshold in the journalism sample, 225 verbs that crossed the threshold in the fiction sample, and 172 such verbs in the scientific-technical sample. The grammatical profiles of the verbs in each sample were fed into a correspondence analysis, which treats each grammatical profile as a vector of numbers (a row with the relative frequencies of the forms) and then calculates the distances between the rows by constructing a multidimensional space defined by mathematically constructed dimensions called "Factors". These Factors are arranged according to their strength in accounting for the variance in the data, such that Factor 1 is the mathematically constructed dimension that is most powerful in sorting the data (in this case, verbs) into two groups: verbs with a positive value for Factor 1 vs. verbs with a negative value for Factor 1. The main finding of this study is that Factor 1 turns out to be interpretable as aspect: Factor 1 consistently sorts the verbs according to aspect, with about 93% accuracy. In other words, given only the

grammatical profile of a verb (which is the only information that the correspondence analysis has access to), it is possible to distinguish perfective verbs from imperfective verbs. Remarkably, the accuracy of this prediction of aspect via grammatical profiles is statistically indistinguishable from the accuracy of prediction via aspectual morphology (prefixes and suffixes). Both grammatical profiles and morphology predict aspect with over 90% accuracy and a chi-squared test comparing the two means of prediction gives a p-value of 0.95, meaning that there is at least a 95% chance that there is no difference between the two.

However, prediction of aspect from grammatical profiles is not fool-proof. A small number of verbs in each of the three samples get misclassified: a few imperfective verbs get wrongly classified as perfectives, and a few perfectives get wrongly classified as imperfectives. This is always due to some strong preference of a verb for a form that is more typical of the opposite aspect. For example, in the fiction sample the imperfective verb *продолжать* 'continue' patterns with perfective verbs because of its strong affinity for the past tense: 76.1% of its attestations in that sample are past tense forms. There is only one verb that is consistently misclassified across all three samples: *смочь* always patterns with the imperfective verbs. In every sample, this deviation of *смочь* is motivated by the fact that it is very frequent in the nonpast, and high relative frequency of nonpast forms is otherwise characteristic of imperfective verbs. In all three samples, *смочь* appears only in indicative forms (no imperatives, infinitives, gerunds, or participles), with the following breakdown, which we can also compare to the numbers for the whole RNC cited above in Section 2.1:

- Journalism: 63% indicative nonpast vs. 37% indicative past
- Fiction: 56.4% indicative nonpast vs. 43.6% indicative past
- Scientific-Technical: 58.8% indicative nonpast vs. 41.2% indicative past
- Whole RNC: 53.12% indicative nonpast vs. 44.62% indicative past

In terms of its grammatical profiles and how they align with aspect, *смочь* is arguably the most deviant verb in Russian. It seems to be masquerading as an imperfective verb, or at least not behaving like a typical perfective verb. This data lends support to Choi's (1994) claims that *смочь* is not the perfective partner of *мочь*.

2.3. Native Speaker Reactions to Use of *смочь* vs. *мочь* Compared with Other Paired Verbs

If *смочь* does not truly function as a perfective partner verb of *мочь*, how do native speakers react to the choice of forms of these two verbs in context? Is

the aspectual distinction clear enough so that native speakers make categorical decisions about their use, or are they to some extent interchangeable? We saw in Section 2.1 that Padučeva (2001) found that past tense forms of these verbs showed a stronger preference for *мог* as opposed to *смог* in Russian two centuries ago than today, and this observation is corroborated by longitudinal data from the RNC. If *смог* is indeed gradually replacing *мог*, can we find evidence for this in the behavior of native speakers?

Janda and Reynolds (under submission) conducted an experiment in which over 500 native speakers of Russian logged their reactions to aspectual choices for verbs in extended authentic contexts. Each participant was randomly assigned to one of six texts of approximately 1100-1600 words each. Participants read the whole text, so all test items were presented in the complete context of the entire text (not just individual sentences). Each test item pair involved a verb for which both a perfective and an imperfective form are morphologically possible, and participants rated both the perfective form and the corresponding imperfective form as "Impossible" = 0, "Acceptable" = 1, or "Excellent" = 2. Participants did not know what the aspect of the verb was in the original text. There was a total of 673 test item pairs in the experiment.

One of the texts contained four sentences with test items relevant to this article, cited in examples (4–7.) This text is an unedited transcript of a guided oral narration videotaped in 2014 at the Multimodal Communication and Cognition Laboratory at Moscow State Linguistic University (MSLU), used by permission from Alan Cienki and Olga Iriskhanova.

(4) В принципе, я могу рассказать об одном случае, когда я не [**смог** / **мог**] уснуть.

'For example, I can tell a story about a situation when I **could**n't fall asleep.'

(5) Я не [**смог** / **мог**] уснуть, потому что примерно два-три года назад у меня ночью была жуткая аллергия, жуткий приступ кашля, и я постоянно кашлял , я не могу уснуть и это происходило вечность.

'I **could**n't fall asleep because about two or three years ago I got an acute allergic reaction in the night, a terrible coughing fit, and I was coughing constantly and I can't fall asleep and it lasted for a long time.'

(6) Я [**смог** / **мог**] быть свидетелем этого.

'I **was able** to witness that.'

(7) Поскольку я опоздал на электричку, был двухчасовой перерыв и я наблюдал за всем этим непосредственно в непосредственной близости и все [**смог** / **мог**] это видеть.

'Since I was late for the commuter train, there was a two-hour wait and I witnessed all that up close and **was able** to see it all.'

The test items are presented in (4–7) in square brackets, and the task was to rate the acceptability of both *смог* and *мог*. In the original versions of all four sentences, the form was *мог* (but this information was not available to participants). Seventy-eight participants completed the ratings for the MSLU text, and their ratings are tallied in Table 2, where the ratings for the non-original form, which is for these test items *смог*, are in shaded boxes and a weighted average is calculated over all the ratings for each item.

Table 2. Ratings of *смог* and *мог* by native speakers in contexts where *мог* is the originally attested form. Ratings of *отлично* scored 2 points, *допустимо* scored 1 point, and *невозможно* scored 0. These numerical ratings are used to calculate the weighted average. The ratings reflect the acceptability of the two forms in sentences (4–7).

Context	Form	*отлично* = 2	*допустимо* = 1	*невозможно* = 0	weighted average
(4)	*смог*	19	45	14	1.06
	мог	63	14	1	1.79
(5)	*смог*	15	46	17	0.97
	мог	65	13	0	1.83
(6)	*смог*	20	16	42	0.72
	мог	48	17	13	1.45
(7)	*смог*	36	24	18	1.23
	мог	54	17	7	1.60

The top two rows of Table 2 can be read as follows. These two rows pertain to the test item pair from the sentence in (4). In the top row, we see ratings for *смог* in sentence (4), where 19 participants rated it as "excellent", 45 rated it as "acceptable", and 14 rated it as "impossible". When these ratings are converted to numerical scores, they yield the weighted average of 1.06 = ((19*2)+45)/78. For the same sentence (4), *мог* was rated "excellent" by 63 participants, "acceptable" by 14 participants, and "impossible" by 1 participant, yielding a weighted average of 1.79 = ((63*2)+14)/78.

We see that in all four sentences the native speakers rated *мог* (which also happened to be the form in the original sentence) more highly than *смог*. However, *смог* also enjoys fairly high ratings and is rated as "excellent" by 15–36 participants. When we compare these results with those for all the other verbs in our experiment, we see an unusually high degree of equivocation for *смог* and *мог*. In our experiment overall, 83% of test pairs received relatively categorical ratings, meaning that one verb form has a weighted average of 1.0 or higher and the other form (of the opposite aspect) has a weighted average of less than 1.0. The *смог*/*мог* test items in (4)-(7) on the contrary receive high ratings for both forms. For two sentences, (4) and (7), both forms received a weighted average over 1.0, and the weighted average of *смог* in sentence (5) is very close to 1. Only the rating of forms in sentence (6) resembles that for the majority of test pairs in our experiment, and even here the results are rather equivocal, since the distance between the two ratings is less than 1. In other words, native speakers seem to find both forms *смог* and *мог* acceptable in this set of sentences, and this level of acceptability is somewhat unusual, since in most contexts native speakers have rather strong preferences for one aspect over the other. In effect, *смог* and *мог* seem to be more similar and interchangeable than other aspectually related pairs of verb forms.

3. Language-Family Evidence: Czech Translation Equivalents

To gain some perspective on the behavior of Russian *смочь*, it could be useful to compare *смочь* with another Slavic language that has inherited the same lexical item. Czech is perhaps the most ideal comparison because it has the etymological equivalent verb, *zmoci* 'achieve', and because the aspect system of Czech provides a contrast as well (cf. Dickey 2000, who finds that Russian and Czech are on opposite ends of the spectrum of Slavic aspectual types). This comparison will give us evidence about the extent to which Russian *смочь* expresses futurity as well as the extent to which the Czech cognate *zmoci* inhabits the same conceptual space as *смочь*.

The ParaSol corpus contains 410 relevant Russian-Czech translation equivalents, 388 obtained by querying for forms of Russian *смочь*, and 22 obtained by querying for Czech *(ne)zmoci*.[6] This data is visualized in Table 3 and Figure 5.

[6] Because Czech orthography requires that the negation be written together with the verb, it was necessary to query for both *zmoci*/*moci* and *nezmoci*/*nemoci* forms.

Table 3. Results of queries for Russian *смочь* and Czech *(ne)zmoci* in the ParaSol corpus. Shaded boxes contain data that is further disaggregated in Figure 5.

	Czech perfective *(ne)zmoci*	**Czech imperfective *(ne)moci***	**Other Czech equivalents**	**Totals**
Russian perfective nonpast *смогу*, etc.	0	79	87	166
Russian perfective past *смог*, etc.	0	91	131	222
Russian imperfective *мочь*	2	NA	NA	2
Other Russian equivalents	20	NA	NA	20
Totals	22	170	218	410

The most frequent translation equivalent of Russian *смочь* in Czech is a form of the imperfective verb *(ne)moci*. Other Czech verbs or phrases that appear often include forms of *(ne)dokázat* '(not) manage', *(ne)umět* '(not) know how (to)', *(ne)podařit se* '(not) succeed', *(ne)být schopen* '(not) be capable', *(ne)být s to* '(not) have the capacity', as well as sentences in which the modality is not overtly expressed. All of these alternatives to *(ne)moci* are represented by the column marked "Other Czech equivalents" in Table 3. The data on translation equivalents of Czech *(ne)zmoci* is scanty, with nearly half of the examples showing no specific equivalent, and the only items appearing more than once are Russian *бессилен* 'helpless', *мочь* 'be able', and *решиться* 'decide'.

In the first row of Table 3 we see that no Russian nonpast forms of *смочь* have Czech equivalents of *(ne)zmoci*. Instead, 79 examples of the Russian nonpast forms appear in Czech as imperfective forms of the verb *(ne)moci* '(not) be able', and a further 87 examples show other translation equivalents in Czech. In the second row which displays equivalents for Russian past forms of *смочь*, again we see no equivalents of *(ne)zmoci*, but 91 equivalents using Czech imperfective *(ne)moci*, along with 131 other equivalents. The next two rows of the table show the Russian equivalents for Czech perfective *(ne)zmoci*, two of which are rendered by forms of Russian *мочь*, while the remaining 20 have other equivalents. The other cells in these rows contain "NA" because the queries were only for Russian *смочь* and Czech *(ne)zmoci* (no queries were conducted for Russian *мочь* or Czech *(ne)moci* or for any other forms).

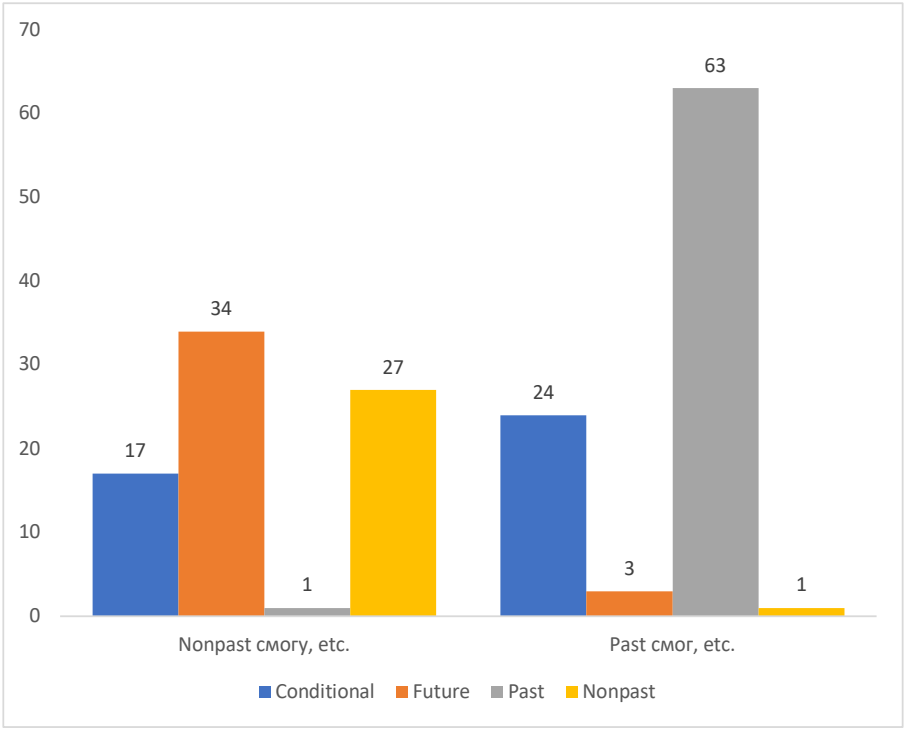

Figure 5. Breakdown of distribution of Czech *(ne)moci* as translation equivalent of Russian nonpast *смогу*, etc. and Russian *смог*, etc. across Czech indicative conditional, future, past, and nonpast

Figure 5 gives a breakdown of the data in the shaded cells of Table 3, where Russian *смочь* is aligned with the Czech imperfective verb *(ne)moci* as its translation equivalent. We see that the equivalents for Russian *смочь* nonpast and past forms include all four types of indicative forms of Czech *(ne)moci*: conditional, future, past, and nonpast. Examples (8)-(12) illustrate those types that appear more than 3 times.

(8) Russian nonpast form of *смочь* parallel to conditional form of Czech *(ne)moci*:
Translations from Umberto Eco. *Il nome della rosa*. 1980

Иначе каждый **сможет** вызывать видения и дурить людей зельями. [*Имя розы*. Елена Костюкович]

jinak by lehkomyslné osoby **mohly** chodit po světě a hlásat lidem svá vidění, neboli lhát s pomocí bylin. [*Jméno růže*. 1985. Zdeněk Frýbort]

(8) ‘otherwise anyone **could** go around announcing their visions and confusing people using herbs.’

(9) Russian nonpast form of *смочь* parallel to future form of Czech *(ne) moci*:
Translations from Stanisław Lem. *Pamiętnik znaleziony w wannie.* 1961.

В будущем я сделаю для вас, что **смогу**. [*Дневник, найденный в ванне.* 1994. К. Душенко.]

V budoucnu udělám, co **budu moci**, samozřejmě služebním postupem. [*Deník nalezený ve vaně.* 1999. Pavel Weigel.]

‘In the future I will do everything **I can**, of course with professional detachment.’

(10) Russian nonpast form of *смочь* parallel to nonpast form of Czech *(ne) moci*:
Translations from Stanisław Lem. *Pokój na Ziemi.* 1987.

Дело в том, что никто, включая и вас, не **сможет** установить, лгут они или говорят правду. [*Мир на земле.* 1990. Е. Невякин.]

Jde o to, *že* nikdo, ani vy sám, ne**může** říct, jestli lžou, nebo jestli mluví pravdu. [*Mír na zemi.* 1989. Helena Stachová.]

‘The point is that nobody, not even you, **can** tell whether they are lying or telling the truth.’

(11) Russian past form of *смочь* + *бы* parallel to conditional form of Czech *(ne)moci*:
Translations from Stanisław Lem. *Fiasko.* 1987.

Я жесток , когда надо быть жестоким, в противном случае тоже не **смог** бы есть мяса. [*Фиаско.* 1991. К. Душенко.]

Jsem bezohledný, když je třeba být bezohledný, jinak bych mimo jiné ne**mohl** jíst maso. [*Fiasko.* 1990. Pavel Weigel.]

‘I am cruel when it is necessary to be cruel, otherwise I **would**n’t be able to eat meat.’

(12) Russian past form of *смочь* parallel to past form of Czech *(ne)moci*:
Translations from Stanisław Lem. *Pamiętnik znaleziony w wannie.* 1961.

Я попытался приподняться, хотя бы выпрямиться, но не **смог** и только повторил… [*Дневник, найденный в ванне.* 1994. К. Душенко.]

Pokoušel jsem se vstát, trochu se narovnat, ale ne**mohl** jsem, jen jsem opakoval… [*Deník nalezený ve vaně.* 1999. Pavel Weigel.]

(12) 'I tried to stand up, to stretch out a bit, but I **could**n't and I just repeated…'

The ParaSol data makes it abundantly clear that Czech *(ne)zmoci* is unlikely to share the semantics of Russian *смочь*, although it has some association with Russian *мочь*, thus bringing more clarity to Barentsen's (2002) suggestion that Russian *смочь* tends to have imperfective translation equivalents in other Slavic languages. However, we find only partial support for Choi's (1994) claim that nonpast forms of *смочь* primarily express futurity in order to compensate for the lack of **буду мочь* in Russian. Although *budu moci* is perfectly grammatical in Czech, as we see in example (9), and although this type of future is the most common single translation equivalent for nonpast forms of *смочь*, the majority of Czech parallels do not use the future, using mostly conditional and nonpast forms of *(ne)moci* instead, as in examples (8) and (10). These examples show that Russian *смочь* is often used in the nonpast without reference to any specific time at all, in what could be called a "gnomic" sense.

4. Language-External Evidence: Spanish Translation Equivalents

Spanish can give us an even more distant perspective on Russian *смочь*. Although both Russian and Spanish of course belong to the same Indo-European language family, they are only distantly related and there are no etymological cognates of Russian *смочь* that could translate that verb. Spanish has an aspectual distinction in the past tense, with the indicative imperfect in some ways similar to the Russian imperfective (and translated as imperfective past in 66.9% of cases), and the indicative preterite similar to the Russian perfective (and translated as perfective past in 85.8% of cases).[7]

The RNC has a parallel corpus of Russian and Spanish texts, the great majority of which are translations into Russian from Spanish (for the purposes of our data, it so happens that all of the relevant examples are from Spanish to Russian translations). This corpus contains 154 examples of forms of Russian *смочь* and their original Spanish equivalents. In 33 of these sentences there is no Spanish verb that serves as the parallel to *смочь*, leaving 121 examples for analysis: 60 of these involve nonpast forms of *смочь*, 60 involve past forms of *смочь*, and one contains the infinitive form *смочь* (which corresponds to the Spanish infinitive *poder* 'be able'). Spanish *poder(se)* 'be able' is by far the most common verb equivalent (85 examples = 70.25%), alongside other verbs such as *lograr* 'manage (to)', *saber* 'know (how to)', *alcanzar(se)* 'achieve', and *conseguir*

[7] These percentages are from Janda and Fábregas forthcoming, a study of verb correspondences in a comparison of the Spanish original of *La Sombra del Viento* by Carlos Ruis Zafón with its Russian translation *Тень ветра*. This data does not include examples where the Spanish original does not correspond to any verb in Russian.

'get'. Table 4 is a confusion matrix of the subparadigms of the original Spanish verb forms that correspond to the 120 examples of nonpast and past forms of Russian *смочь* in this corpus. Examples (13)-(16) illustrate the most common uses of Russian *смочь* to translate forms of Spanish *poder* boldfaced in Table 4.

Table 4. Subparadigms of Spanish verbs translated as Russian *смочь*

Spanish verb forms	**Nonpast *смогу*, etc.**	**Past *смог*, etc.**
Conditional	**11**	6
Future	**14**	1
Imperfect	5	0
Present	**10**	1
Preterite	2	**43**
Infinitive	2	0
Perfect (Present, Past, and Subjunctive Past)	2	6
Subjunctive Imperfect	6	2
Subjunctive Present	8	1

(13) Spanish *poder* conditional translated as nonpast of *смочь*

No **podría**, todo me huele a cebolla. [Camilo José Cela. La Colmena (1951)]

Я не **смогу**, мне все пахнет луком.

'I **can**'t, everything smells like onion to me.'

(14) Spanish *poder* future translated as nonpast of *смочь*

Mi pobre hijo, que se está poniendo muy delicado de salud, no **podrá** trabajar. [Benito Pérez Galdós. Doña Perfecta (1876)]

Бедный мальчик в последнее время так ослабел, что скоро совсем не **сможет** работать.

'The poor boy has gotten so weak of late that soon he **wo**n't be able to work at all.'

(15) Spanish *poder* present translated as nonpast of *смочь*

Pues me lo dice y yo, si **puedo**, se lo arreglo. [Camilo José Cela. La Colmena (1951)]

Скажите мне, и я, если **смогу**, помогу вам.

'Just tell me, and if I can, it will be arranged.'

(16) Spanish *poder* preterite translated as past of *смочь*

Los mandos eran cargos políticos y el solo un jefe inferior, así que no **pudo** hacer nada. [Manuel P. Villatoro. «El corsario español Antonio Barceló machacó el nido de piratas de Argel sin tener apenas bajas» *[www.abc.es] (2016.12.20)]*

Операцией руководили высокие чины, а он был всего лишь местным командиром, поэтому ничего не **смог** сделать.

'The commanders were high-ranking politicians, and he was just a junior officer, so he **could**n't do anything.'

The frequencies in the right-hand column of Table 4 show that the majority of Spanish preterite forms (43) are translated as Russian past tense forms of *смочь*, suggesting that the past tense forms of *смочь* do indeed behave like perfectives according to this measure. As concerns the nonpast forms of *смочь*, this distribution very much resembles the distribution of Czech translation equivalents with forms of *(ne)moci*: the largest number of forms correspond to future tense (34 forms, which is 43% of the Czech data for translation equivalents of nonpast forms of *смочь*), but future does not make up a majority and is outweighed by the combination of conditional and present tense forms (17 + 27 = 44 forms in Czech). Again, we find only weak support for Choi's hypothesis concerning the expression of futurity by nonpast forms of *смочь*; however this comparison is compromised by the fact that the datasets are small and reflect different directions of translation.

5. Conclusion

I have presented a variety of corpus and experimental data documenting the behavior of Russian *смочь* in comparison with *мочь*, with other Russian verbs, with Czech cognates, and with Spanish translation equivalents. There is no evidence that Russian *смочь* shares any semantic overlap with Czech *(ne)zmoci*. There is ample evidence that *смочь* is a peculiar verb, particularly when we compare its grammatical profiles to those of other Russian verbs: nearly all of its forms attested in corpora are indicative, and nonpast forms comprise the majority, despite the fact that past tense forms normally predominate for perfective verbs. Native speakers are more equivocal in their rating of the

acceptability of past tense *смочь* in contexts where past tense *мочь* is used, when compared with acceptability ratings of other aspectually paired Russian verbs, where the results tend to be more categorical. However, past tense forms of *смочь* behave very similarly to other past perfective forms that serve as translation equivalents of Spanish preterites. We also find that *смочь* has become significantly more frequent in Russian over the past century, but that it is not possible to connect this rise in frequency directly to an expression of futurity that would make up for the lack of forms like **буду мочь*.

References

Barentsen, Adriaan. (2002) "O russkix glagolax *smoč'* i *sumet'*". V.V. Ivanickij i dr., eds. *Osnovnye problemy russkoj aspektologii*. Sankt-Peterburg: Nauka, 7–29.

Choi, Sung-ho. (1994) *Modal predicates in Russian: Semantics and syntax*. Ph.D. Dissertation, University of California, Los Angeles.

———. (1999) "Semantics and syntax of *мочь* and *смочь*: Their 'aspectual' relationship". *Russian linguistics* 23: 41–66.

Dickey, Stephen M. (2000) *Parameters of Slavic aspect: A cognitive approach*. Stanford, CA: CSLI Publications.

Divjak, Dagmar. (2010). *Structuring the lexicon: A clustered model for near-synonymy*. Berlin: De Gruyter.

Eckhoff, Hanne M., Laura A. Janda, and Olga Lyashevskaya. (2017) "Predicting Russian aspect by frequency across genres". *Slavic and East European journal* 64(1): 844–75.

Janda, Laura A. and Dagmar Divjak. (2008) "Ways of attenuating agency in Russian". Anna Siewierska, ed. *Impersonal constructions in grammatical theory = Transactions of the Philological Society* (v. 106), 138–179.

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): 719–63.

Janda, Laura A. and Robert J. Reynolds. Forthcoming. "Construal vs. redundancy: Russian aspect in context".

Janda, Laura A. and Antonio Fábregas. Forthcoming "Seeing from without, seeing from within: Aspectual differences between Spanish and Russian".

Padučeva, Elena. V. (2001) "Russkij literaturnyj jazyk do i posle Puškina". G. Ressel, ed. *A. S. Puškin und die kulturelle Identität Russlands*. Frankfurt am Main: Peter Lang, 97–108.

v. Waldenfels, Ruprecht. (2011) "Recent developments in ParaSol: Breadth for depth and XSLT based web concordancing with CWB". M. Daniela and

R. Garabík, eds. *Natural language processing, multilinguality*. Proceedings of Slovko 2011, Modra, Slovakia, 20–21 October 2011. Bratislava, 156–62.

UiT, The Arctic University of Norway
laura.janda@uit.no