**Construal vs. Redundancy: Russian Aspect in Context**

**Abstract**

The relationship between construal and redundancy has not been previously explored by empirical means. We investigate the relationship between redundancy and construal by undertaking an experiment. Russian aspect offers speakers the opportunity to construe situations as either Perfective or Imperfective. In our experiment, 500 native Russian speakers rated the acceptability of both Perfective and Imperfective verb forms in extensive authentic contexts representing various genres. We find that the meanings of Perfective and Imperfective are firmly anchored in about 81% of uses, where the choice of aspect is largely redundant given the context. In these redundant uses, native speakers, on the whole, give higher ratings to the original aspect than to the opposite aspect. By contrast, in about 17% of contexts, aspect is open to relatively free interpretation by the speaker. We contend that it is the anchoring in redundant contexts that likely facilitates the independence of construal in contexts with less redundancy. We also find that native speakers differ in their reaction to stimuli, and this variation is greatest when the stimuli have been artificially manipulated. Native speakers are fairly consistent in giving the original aspect high ratings, but much less consistent in rating the non-original aspect.

**1. Introduction: Construal, Redundancy, and Russian Aspect**

Cognitive linguists readily acknowledge the pervasiveness of construal and redundancy in language. However, little attention has been paid to the relationship between construal and redundancy. We argue that redundancy actually facilitates construal, and back this argument up with an empirical study of the relationship between redundancy and construal. Our usage-based approach shows how construal and redundancy co-exist on a scale, and how redundancy anchors the meanings available for construal. In this section, we review the concepts of construal and redundancy, as well as the distinction between Perfective and Imperfective aspect in Russian, which serves as the topic of our investigation. Section 2 presents our experiment, the stimuli it contained, and how it was carried out. In Section 3 we first check to see whether frequency has an effect on the rating of Perfective vs. Imperfective verb forms. We find that frequency does not have a significant effect on our data, and we analyze our results in terms of what they mean not only for Russian aspect, but also for redundancy and construal, as well as variation across speakers. We present our conclusions and suggest directions for future research in Section 4.

**1.1 Construal**

In reporting the same event, I can say either *A pickpocket stole Frank’s money* or *Frank’s money got stolen by a pickpocket.* Active vs. Passive voice in English (and many other languages) provides me as a speaker with the opportunity either to focus on the Agent, putting the pickpocket in Subject position, or to take the perspective of the victim, putting the Patient, Frank’s money, in Subject position. Voice is just one of a multitude of grammatical distinctions that facilitate alternative construals of meaning in language. Langacker’s (2015: 120) most recent definition of construal states simply: “*Construal* is our ability to conceive and portray the same situation in alternate ways”. This definition emerges from a persistent focus within cognitive linguistics on the relationship between perception and conception, which Verhagen (2007: 50) traces back at least as far as Talmy (1978), who also coined the term *ception* (Talmy 1996) to emphasize the indissoluble relationship between the two cognitive mechanisms. Language does not directly encode reality. Instead, all linguistic expressions contain an element of construal (Croft & Cruse 2004: 69, Verhagen 2007: 48), which can be thought of as a way of *viewing* a situation (Langacker 2008: 261). It is no accident that the Russian term for aspect, *vid* (literally ‘type’), is transparently a deverbal noun derived from *videt’* ‘see’. In Russian, aspect is quite literally a “viewing” of an event.

Construal plays a crucial role in meaning, which “consists of both conceptual content and a particular way of construing that content” (Langacker 2008: 4). Two expressions may refer to the same event, but this does not entail that they have the same meaning, because although they have the same conceptual content, the meanings of the two expressions differ because of differences in construal (Langacker 2008: 95).

Although many discussions of construal have focused on distinctions available in the lexicon (such as *steal* which focuses on the item taken, vs. *rob* which focuses on the victim), it is acknowledged that construal is equally important for grammatical distinctions (see numerous examples in Croft & Cruse 2004: Chapter 3, Langacker 2008: Chapter 3, Langacker 2015: 121). We argue that verbal aspect is an excellent example of a grammatical distinction that offers speakers alternate ways to construe events.

There have been several attempts to classify types of construal across various parameters, such as salience, scope, and perspective (cf. Verhagen’s 2007 survey focusing primarily on works by Langacker and Croft & Cruse), without yielding an agreed-upon classification. Langacker (2015: 121) finds the task too unwieldy: “Construal encompasses numerous interrelated factors. While natural groupings can be observed, no one classificatory scheme captures all the relationships or does justice to a single factor.” Russian aspect, as explained below, likewise makes reference to a variety of factors in the construal of events, including salience, scope, and perspective.

Both Taylor (2002: 11, 386) and Verhagen (2007: 49) observe that languages differ in the options for construal that they provide for their speakers. This observation echoes Jakobson’s (1959/1971: 492) famous statement 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”, since each language requires its speakers to make different set of choices among alternate construals. Plungian (2000: 109) likens these choices to a “grammatical questionnaire” that speakers must continuously fill out, and notes that mastering this task is one of the second language learner’s greatest challenges. Langacker (2008: 284-296) gives the example of English definiteness as a choice continuously imposed upon English speakers, but absent in many other languages. Russian aspect is similarly ubiquitous, requiring speakers to express Perfective or Imperfective with every verb form.

Croft & Cruse (2004: 75-103) present construal as a dynamic process that recruits information from the linguistic context, the physical and social context, as well as encyclopedic and stored knowledge about one’s language. All of this information guides the use of construal to provide meaning. Context is, however, messy territory, and there has been little exploration of the contexts that support alternative construals. A notable exception is, for example, Bresnan’s (2007) exploration of the *to*-dative/ditransitive alternation in English between *John gave a book to Sally* (*to*-dative) and *John gave Sally a book* (ditransitive), where a multitude of factors (use of pronouns, animacy of the recipient, relative length of object phrases) make one or another variant more or less likely, as in *John whispered Sally the answer* or *The chilly weather gave Sally a cold* (compare with …*to Sally* in both). Our study makes it possible to identify concrete contexts that favor Perfective or Imperfective aspect in Russian or even allow both aspects.

**1.2 Redundancy**

Redundancy has received considerably less attention in cognitive linguistics.[[1]](#footnote-1) Dahl (2004: 9-13, 187-188, 291-295) discusses redundancy as a means of preventing information loss, and claims that languages have a tendency to develop system-level “smart redundancy” that makes it possible to maximize the advantages of redundant elements without proliferating them. This is achieved through successive cycles of redundancy-increasing and redundancy-decreasing changes that ultimately increase smart redundancy. Dahl finds that languages with “mature features” (those associated with linguistic complexity) make the most use of smart redundancy. Langacker (2008: 188-189) identifies redundancy as a pervasive phenomenon in language and associates it more narrowly with grammatical agreement. According to both Dahl and Langacker, redundancy is neither superfluous nor meaningless. Instead, redundancy provides the hearer with extra clues which can be helpful in decoding an imperfect message, and provides the speaker with the means to emphasize the same information in a variety of ways. Langacker eschews the term *agreement* on the grounds that the redundant information is not “copied” from one part of an expression to another, but supplied as multiple manifestations in a complex expression. This means that the redundant elements are not semantically empty or superfluous, but instead overlap. However, neither Dahl nor Langacker address the relationship between construal and redundancy.

We can return to Langacker’s example of English definiteness as an example of a construal that is usually redundant in context. The use of English articles bedevils even the most superbly articulate L2 users whose native languages lack definiteness. As any native speaker of English who has edited texts written by these L2 speakers can attest, in most otherwise well-formed sentences it is quite easy for L1 users of English to add in missing articles and correct the ones that are misplaced. This is because the construal that English definiteness conveys is usually recoverable from context. If just one eligible candidate for reference is available in a context, then the reference has to be definite, as, for example, when we use a word like *only* or a superlative like *toughest* (Langacker 2008: 287). In these contexts (and many others), English requires us to use *the*, but the definite article is redundant. It is the very redundancy of articles that likely makes them so hard to master in the first place: Ellis & Wulff (2015: 420) cite numerous studies showing that L2 learners have a hard time mastering grammatical markers that are redundant in their understanding of an utterance.

However, English articles are not always redundant. There are situations where either article can be used depending on what the speaker wants to say. If I say *Bring me the cup*, I am telling the hearer that I have a particular cup in mind, and it is the speaker’s job to figure out which one. But I can also say *Bring me a cup*, without making any assumptions about which cup I should get. There are many similarities between English definiteness and Russian aspect, which is sometimes even termed “temporal definiteness” (Dickey 2000, Dickey & Janda 2015).

Our point in this article is that it is the redundant uses of a distinction that anchor the meanings that are relevant for construal. These redundant uses reinforce the construals associated with the markers. This continuous strengthening of association facilitates the use of those same markers to invoke their associated construals also in contexts where they are not redundant. Our study measures the relationship between redundancy and free construal of Russian aspect, and shows that these are not discrete phenomena, but actually two endpoints of a continuum.

**1.3 Russian aspect**

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), and in other languages aspect is often optional and/or allows for a “neutral” aspect (as in Chinese). In Russian, all forms of the verb obligatorily express either Perfective or Imperfective aspect.[[2]](#footnote-2) 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. A vast scholarly literature (cf. references in Dickey 2000, Zaliznjak and Šmelev 2000, Timberlake 2004, Janda 2007b, Janda et al. 2013) is devoted to Russian aspect, which is not fully mastered even at age six by native speakers (Stoll 2001, Gagarina 2004), and is routinely listed as the single greatest obstacle facing second language learners of Russian (cf. Offord 2005, Andrews et al. 1997, Cubberly 2002, Martelle 2011).

From the perspective of cognitive linguistics, Janda (2004) has described the meanings associated with Russian Perfective and Imperfective in terms of a complex version of a TIME IS SPACE metaphor. This metaphor treats situations as material entities, and more specifically invokes these mappings: PERFECTIVE IS A DISCRETE SOLID OBJECT vs. IMPERFECTIVE IS A FLUID SUBSTANCE. The contrast of two types of physical matter, discrete solid objects vs. fluid substances, provides a rich source domain with a multitude of properties that align with the meanings of the two aspects. For example, discrete solid objects are unique individuals with edges, are rigid, can have various shapes (including very thin/punctual), are perceptually salient (foreground), but cannot flow and cannot be easily penetrated (appear as Gestalts). Imperfective situations, like fluid substances, lack inherent shape and boundaries, but can both flow and spread and fill containers and be easily penetrated (described from the inside), are perceptually diffuse (background), and can be mixed together. This isomorphism between matter and aspect accounts for many facts about the use of Russian aspect, for example that Perfective is used to describe bounded events seen as wholes, that can be of varying duration (even punctual), that are necessarily sequenced (and therefore mostly incompatible with present tense), used to describe prominent plot-line events and to express successful completion and issue instructions and warnings. Imperfective on the other hand is used to express unbounded situations that require at least some duration (never punctual), can express situations that are simultaneous with each other and with the present tense, can be used in gnomic expression of eternal facts, can describe gradual processes and repeated actions, can describe how an action unfolds, and is used to describe background situations (settings), trying to do something, categorical negation, frustration, and polite imperatives in certain social situations. In other words, Russian can construe a situation either as the temporal correlate of a discrete solid object or as the temporal correlate of a fluid substance.

Russian imposes on its speakers the choice between these two alternative construals of situations as Perfective vs. Imperfective every time they use a verb. This choice involves various types of construal, including focusing, scope, and profiling (cf. Langacker 2008: Chapter 3). In terms of focusing, the Perfective gives a foregrounding construal as opposed to the Imperfective, which serves to background events. In terms of scope, the Perfective includes an entire bounded event, as opposed to the Imperfective which has more limited scope, excluding the endpoints of the bounded event. In this sense, Russian Imperfective aspect is in some ways similar to (but by no means equivalent to) the English progressive. To a limited extent, Russian Perfective corresponds to Langacker’s summary scanning as opposed to the Imperfective as sequential scanning, but this correspondence fails to reflect certain language-specific details. In terms of profiling, the Perfective profiles the completion of an event, whereas the Imperfective does not profile the completion.

In many contexts, the choice of Perfective vs. Imperfective aspect is redundant in Russian. Descriptive grammars of Russian list dozens of adverbials and other syntactic “triggers” that (usually) indicate that only one aspect is available. For example, *za tri minuty* ‘in three minutes’ is a trigger for Perfective verb forms (*Ja s”ela banany za tri minuty* ‘I ate the bananas in three minutes’ where *s”ela* ‘ate’ is a Perfective verb form), while *vsegda* ‘always’ is a trigger for imperfective verbs (*Ja vsegda ela banany* ‘I always ate bananas’ where *ela* ‘ate’ is an Imperfective verb form). Some examples of typical triggers are presented in Table 1 (a subset of a compilation taken from the sources listed in Appendix 1).

|  |  |  |
| --- | --- | --- |
|  | **Adverbials as Triggers** | **Verbs as Triggers** |
| **Preference for Perfective verb forms** | *nakonec* ‘finally’,  *vnezapno* ‘suddenly’, *srazu* ‘immediately’, *čut’ ne* ‘nearly’, *vdrug* ‘suddenly’, *uže* ‘already’, *neožidanno* ‘unexpectedly’, *sovsem* ‘completely’, *za tri minuty* ‘in three minutes’… | *zabyt*’ ‘forget’, *ostat’sja* ‘remain’, *rešit*’ ‘decide’, *udat’sja* ‘succeed’, *uspet*’ ‘succeed’, *spešit*’ ‘hurry’… |
| **Preference for Imperfective verb forms** | *vsegda* ‘always’, *často* ‘often’, *inogda* ‘sometimes’, *poka* ‘while’, *postojanno* ‘continually’, *obyčno* ‘usually’, *dolgo* ‘for a long time’, *každyj den*’ ‘every day’, *vse vremja* ‘all the time’, *tri časa* ‘for three hours’…  categorical negation: *ne nado* ‘should not’, *ne stoit* ‘not worth’*, ne razrešaetsja* ‘not allowed’… | Phasal verbs[[3]](#footnote-3): *stat*’ ‘start’, *načat’/načinat’* ‘begin’, *prodolžit’/prodolžat’* ‘continue’, *končit’/končat’* ‘stop’  Verbs of motion: *pojti* ‘go’, etc.  Others: *učit’sja* ‘learn’, *umet*’ ‘know how’, *ljubit*’ ‘love’… |

Table 1: Some examples of “triggers” for Perfective and Imperfective aspect in Russian

Reynolds (2016) tested the behavior of these triggers in aggregate against corpus data. While he found that they indeed predict aspect with fairly good reliability (around 96%), even when taken in aggregate, these triggers are relatively rare in actual language use, appearing in association with only about 2% of verbs in corpus language samples. In other words, descriptive grammars fail to represent 98% of the relationship of context to aspect.

While variable construal of aspect is acknowledged as a significant phenomenon (cf. Langacker 2008: 147-160), there has not been previous experimental investigation of relevant factors. We conducted an experiment in order to reveal contexts where selection of Perfective vs. Imperfective aspect is inflexible in Russian, as opposed to contexts where aspect is open to construal. We find that there are indeed contexts where the construal of aspect is redundant, but there are also contexts where aspect is open to construal, as well as a range of contexts that combine redundancy and openness to construal to varying degrees.

**2. The Experiment**

We conducted an experiment to test the reactions of native speakers of Russian to the choice of Perfective vs. Imperfective aspect. This is, to our knowledge, the first experiment on Russian aspect using full-sized authentic contexts as opposed to single sentences or short passages that are specially constructed for an experiment. In Section 2.1, we describe the stimuli selected for the experiment, in Section 2.2 we present the experimental procedure, and in Section 2.3 we describe the recruitment of participants for the experiment.

Previous experimentation on aspect with Russian native speakers has been more limited or simply different in its scope. The most relevant previous experiment was conducted by Gorbova (2010), who reports on data from forty-five Russian native speakers who are college students studying Romance languages. Participants were shown brief fragments of texts (1-3 sentences) translated from various Spanish works of literature and asked to underline one of three forms (Imperfective Past, Perfective Past, and Imperfective Non-Past) that was most felicitous in the given context (or to rate two or all three options if more than one was possible). It was found that for about half of the test items, native speakers were largely in agreement that only one form was possible, and that over half of these were contexts with no triggers (called “catalyzers” by Gorbova). In the remaining test items, there was lack of agreement among speakers, who often found two or even all three items acceptable. Our experiment differs from Gorbova 2010 in that it uses complete contexts of authentic texts originally written in Russian, as well as a larger number of participants.

Other previous experiments have even less in common with our experiment, usually restricted only to verbs (in the absence of full contexts) or populations with incomplete acquisition of aspect (young children or heritage speakers). A survey of thirty-six linguists (not all of them native speakers) in 1997 (Anketa 1997) asked whether seven pairs of Russian verbs constituted aspectual pairs (where two verbs differ only in aspect), and received very divergent answers (see discussion in Gorbova 2011). Batiukova et al. (2012) report on a semantic decision task, in which participants were shown a verb prime and then a target verb, and had to decide whether the target verb “refers to an event / situation with a clear outcome”. They found that resultative verbs were recognized more quickly, especially if they had prefixes marking the aspect. Vinnitskaya & Wexler (2001) conducted three experiments on children: a narrative task, a comprehension task, and an elicitation task, finding that children use more Imperfective verb forms than adults. Stoll (2001) presents a more extensive series of experiments on children’s acquisition of aspect, and finds that their use of aspect is more conservative than that of adults and that they achieve better mastery of resultative uses for Perfective than for atelic uses. Polinsky (2008) undertook an experiment with heritage speakers of Russian raised in the US. Nine speakers of similar proficiency listened to 90 sentences, all of which contained the combination of a matrix verb followed by an Imperfective or Perfective Infinitive, some of which were felicitous in Russian, and some of which were not. This was a forced-choice experiment in which participants had to choose whether each sentence was grammatical or not. Polinsky found that heritage speakers of Russian exhibit imperfect acquisition of aspect, often using only one aspect per lexical concept.

**2.1 Stimuli**

Six texts served as stimuli for our experiment, as shown in Table 2 (see also sources and dates of origin of these texts in Appendix 2). Explanations of columns 3 and 4 in Table 2 are found in this section; explanations of columns 5 to 7 are found in Section 2.3. The stimuli were carefully chosen in order to meet a series of criteria: authenticity, balance across registers/genres, length, density of test items, appropriateness, and permission for use. Full versions of the stimuli can be accessed at URL [TROLLing post].

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Genre** | **Abbreviated Title** | **Words** | **Items (pairs)** | **Participants** | **Outliers** | **Participants - Outliers** |
| Fiction | Beetle | 1459 | 300 (150) | 83 | 1 | 82 |
| Journalistic Prose | Summit | 1116 | 166 (83) | 84 | 4 | 80 |
| Scientific-Technical Prose | Phages | 1558 | 198 (99) | 72 | 1 | 71 |
| Spoken Narration | Yellow Sign | 1275 | 160 (80) | 99 | 4 | 95 |
| Guided Spoken Narration | MSLU | 1617 | 278 (139) | 78 | 2 | 76 |
| Radio Interview | Ivan D. | 1468 | 244 (122) | 85 | 3 | 82 |
| *Totals* |  |  | *1346*  *(673)* | *501* | *15* | *486* |

Table 2: Overview of Experimental Stimuli

*Authenticity*

Only contemporary authentic texts were chosen. All six texts were created by and for native speakers (none were elicited for this experiment). Five of the texts date from 2015-2016, and one is from 2004. Authenticity of stimuli is essential if we are to understand how Russian aspect interacts with context in real use.

*Balance*

Because we wished to have a balanced representation of Russian and also because we wanted to check whether there might be a difference in aspectual usage across genre[[4]](#footnote-4), half of the texts represent written genres, and half represent spoken genres. Within each register, further distinctions were made. For the written register, we have one text representing Fiction (an entire short story), one representing Journalism (an entire article about the politics of petroleum), and one representing Scientific-Technical Prose (an entire article about scientific experiments conducted on bacteriophages). All texts in the written register were retrieved from internet sources. For the spoken register, we have one Narration (in which the speaker tells an amusing story) taken from a corpus of spoken Russian collected by Andrei A. Kibrik and colleagues, one Guided Narration (in which the speaker gives lengthy responses to questions about his background and experiences he has had) taken from a collection of interviews of college students conducted at Moscow State Linguistic University, and one radio interview of a war veteran taken from the Russian National Corpus. The speech of the war veteran contained some dialectal features. None of the texts were abridged or otherwise edited in any way.

*Length*

The goal was to provide the fullest context possible in an experimental setting so that no information would be absent due to lack of context. A pilot study showed that the longest text that could be accommodated without overtaxing participants was about 1500-1600 words. At this length, most participants in the pilot study were able to complete the task in under twenty minutes. The texts also needed to be approximately the same length. As shown in Table 2, the length of the texts varied from 1116 to 1617 words, with an average of 1415.5 words per text.

*Density of Test Items*

Since one objective was to get as much data as possible on aspectual choices, it was important to select texts that provided plenty of possible choices, avoiding texts with long passages without such choices. Texts were carefully selected to maximize the number of verb pairs that participants would be exposed to.

Russian verbal conjugation has only two tenses, one that is Past and one that is not. The latter tense is often referred to as the “Non-Past”, and in the Non-Past, Imperfective verbs usually express Present tense, while Perfective verbs usually express Future tense. In addition, there is a periphrastic Future used with Imperfective verbs. This distribution is represented in Table 3, illustrated with forms from the verb pairs *napisat’* ‘write [Perfective]’ / *pisat’* ‘write [Imperfective]’ and *vyigrat’* ‘win [Perfective]’ / *vyigryvat’* ‘win [Imperfective]’. In the first verb pair (‘write’) the difference in aspect is marked with the prefix *na-* on the Perfective verb forms, while in the second verb pair (‘win’), the difference in aspect is marked with the suffix -*yva* on the Imperfective verb forms. These two verb pairs represent two of the most common ways in which aspect is marked morphologically in Russian.[[5]](#footnote-5)

|  |  |  |
| --- | --- | --- |
|  | **Perfective** | **Imperfective** |
| **Past** | *napisal* ‘he wrote’  *vyigral* ‘he won’ | *pisal* ‘he wrote’  *vyigryval* ‘he won’ |
| **Present** | [No aspectual contrast available] | *pišet* ‘s/he writes’  *vyigryvaet* ‘s/he wins’  (Non-Past) |
| **Future** | *napišet* ‘s/he will write’  *vyigraet* ‘s/he will win’  (Non-Past) | *budet pisat’* ‘s/he will write’  *budet vyigryvat’* ‘s/he will win’  (periphrastic Future) |
| **Infinitive** | *napisat’* ‘write’  *vyigrat’* ‘win’ | *pisat’* ‘write’  *vyigryvat’* ‘win’ |
| **Imperative** | *napišite* ‘write!’  *vyigrajte* ‘win!’ | *pišite* ‘write!’  *vyigryvajte* ‘win!’ |

Table 3: Aspectual contrasts available in Russian

Table 3 shows how Perfective vs. Imperfective aspect contrast in four subsets of the verbal paradigm, namely in all Past, Future, Infinitive, and Imperative forms. The test items targeted precisely these four grammatical categories, excluding all verb forms where an aspectual contrast was not available.

The following types of verb forms were not rendered as test items on the grounds that they precluded aspectual contrast:

* Present tense, since only Imperfective verbs can express present tense;
* Forms of verb *byt*’ ‘be’, since this verb is always Imperfective in Russian;
* Gerunds and participles, which tend to be highly specific to one aspect or the other (Present gerunds and Present participles are largely restricted to Imperfective, while Past gerunds and Past participles are largely restricted to Perfective aspect);
* Bi-aspectual verbs which express both Perfective and Imperfective aspect without any morphological distinction, such as *rodit’sja* ‘be born’, *realizovat’* ‘realize, implement’;
* Other verbs not paired for aspect, such as -*sja* passives like *prednaznačat’sja* ‘be intended for’ and the verb *stat’* when used as a phasal verb meaning ‘begin’ (though *stat’* could serve as a test item when it appeared as a main verb meaning ‘become’)

When these types of verbs appeared in the stimuli, they were merely rendered as is, rather than being presented as test items. In other words, they were just part of the rest of the text that the participants read.

Example (1) demonstrates how original verb forms were rendered as pairs of test items in the opening lines of the ‘Beetle’ text:

(1)(a) Original text:

*Prav-o vybor-a žiznennogo put-i -- bol’š-oj podarok sud’b-y. U Vasilij-a èt-ogo prav-a ne* ***by-l-o****. On bezropotno* ***prinja-l*** *vybor, kotor-yj za nego* ***sdela-l-a*** *sud’b-a, i èt-o* ***by-l*** *velik-ij šag.*

[right-Nom.Sg choice-Gen.Sg life-Gen.sg path-Gen.Sg big-Nom.Sg gift.Nom.Sg fate-Gen.Sg at Vasilij-Gen.Sg that-Gen.Sg right-Gen.Sg not **be-Past-Neut.Sg** he.Nom.Sg uncomplainingly **accept-Past.Masc.Sg** choice.Acc.Sg which-Acc.Sg.Masc for he.Acc.Sg **make-Past-Fem.Sg** fate-Nom.Sg and that-Neut.Sg **be-Past.Masc.Sg** big-Nom.Masc.Sg step.Nom.Sg]

‘The right to choose one’s path in life is a great gift of fate. Vasilij **didn’t have** that right. He uncomplainingly **accepted** the choice that fate **made** for him, and that **was** a major step.’

(1)(b) Text with test item pairs:

*Pravo vybora žiznennogo puti -- bol’šoj podarok sud’by. U Vasilija ètogo prava ne bylo. On bezropotno* ***[prinjal / prinimal ]*** *vybor, kotoryj za nego* ***[sdelala / delala ]*** *sud’ba, i èto byl velikij šag.*

The example contains four overt verb forms boldfaced in (1)(a): *bylo* ‘was’ (a Past form of Imperfective *byt*’ ‘be’), *prinjal* ‘accepted’ (a Past form of Perfective *prinjat’* ‘accept’), *sdelala* ‘made’ (a Past form of Perfective *sdelat’* ‘make’), and *byl* ‘was’ (another Past form of Imperfective *byt*’ ‘be’). In (1)(b) the two forms of *byt*’ ‘be’ were left as is since there is no aspectual contrast available, and the other two verb forms are presented as test item pairs (boldfaced), including the corresponding Imperfective forms *prinimal* ‘accepted’ (a Past form of Imperfective *prinimat’* ‘accept’) and *delala* ‘made’ (a Past form of Imperfective *delat’* ‘make’). The purpose was to collect ratings from participants of both the Perfective and the Imperfective verb forms for every pair, without providing any indication of what the original aspect was.

The fourth column in Table 2 indicates the number of test items in each text. Each verb form that did allow for aspectual contrast was rendered as two test items by offering the participants verb forms for both Perfective and Imperfective aspect. Table 2 shows both the number of test items and the number of verb pairs (given in parentheses) in each stimulus text. For example, the ‘Beetle’ text contains 300 test items, which are 150 Perfective/Imperfective verb pairs based on original verb forms of either aspect. The density of test items per text can be calculated based on the figures in columns 3 and 4 in Table 2. The ‘Beetle’ text is the most dense, with 10% of words yielding pairs of test items, while the ‘Yellow Sign’ text is the least dense, with 6% of words yielding test pairs. However, the densities of the six stimuli are very similar, since texts with lower densities were not selected.

One consequence of using authentic texts and the other criteria we had was that we had little control over the distribution of test items across subparadigms. Table 4 shows how many test item pairs there were for various original combinations of subparadigm and aspect. For example, the first row of Table 4 indicates that there were 298 Perfective Past tense verb forms that served as the basis for test item pairs, and 134 Imperfective Past tense verb forms that served as the basis for test item pairs. The distribution in Table 4 reflects the same kind of distribution discovered by Janda & Lyashevskaya (2011) in their investigation of nearly six million verb forms extracted from the Russian National Corpus: there are more Perfective forms than Imperfective forms overall, and for both aspects there are more Past forms than Infinitives, and more Infinitives than Imperatives.[[6]](#footnote-6) As a result of this general skew in the distribution of verb forms in Russian, we have most of our data on original uses of the Past tense (both Perfective and Imperfective) and the Perfective Infinitive. We also have quite robust data for the Perfective Future and Imperfective Infinitive, but progressively less data for the Perfective Imperative, Imperfective Imperative, and Imperfective Future. However, this skewed distribution also contributes positively to the naturalness of the stimuli since it follows the distribution that is typical for Russian.

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Perfective** | **Imperfective** | ***Total*** |
| **Past** | 298 | 134 | *432* |
| **Future** | 50 | 7 | *57* |
| **Infinitive** | 104 | 60 | *164* |
| **Imperative** | 10 | 10 | *20* |
| ***Total*** | *462* | *211* | *673* |

Table 4: Number of test item pairs for each original combination of aspect and subparadigm

*Appropriateness and Permission*

Stimuli were carefully chosen to avoid taboo language and subjects, as well as politically sensitive or potentially unpleasant topics. We sought and obtained the rights to use all of the texts that were included in the experiment.

**2.2 Experimental Procedure**

The experiment was conducted as an online survey using Qualtrics software[[7]](#footnote-7). Each participant completed a “CAPTCHA” task to prove that they were human (not a robot) and was then randomly assigned to one of the six texts that served as stimuli. The instructions told the participants that they would find in square brackets pairs of verbs for them to rate according to their acceptability in the context of the text as “Excellent”, “Acceptable”, or “Impossible”, that the choice was to be made according to their personal opinion, and that it was possible to give both verbs the same rating. They were also told that they must complete the entire experiment, rating all of the verbs in square brackets, in order to receive a code for the lottery (see Section 3.3). They were warned that their code would be eliminated from the lottery if there was evidence that they just filled in the survey at random without reading the text, and were told that each participant could fill in only one survey. Participants were also told that their participation was voluntary and they could quit the experiment at any time, and that by participating they were giving their consent and agreeing to the terms of the experiment. Prior to reading the text and evaluating the verbs, participants were asked to state their age (to confirm that they were 16 or older), their native language (to confirm that they were native speakers of Russian), and their gender identity. No IP addresses or other identifying information was collected.



Figure 1: Screenshot illustrating what the experiment looked like

Figure 1 shows the beginning of the survey for the ‘Beetle’ text. The instructions are in blue and the survey text follows the horizontal line. The test pairs are in square brackets and highlighted in light blue, with the Perfective verb form followed by a slash and then the corresponding Imperfective verb form. There is no indication of what the original verb form was in the source text. When the participant moves the cursor over a verb in a test pair, the three evaluations pop up as illustrated in Figure 1. Here the participant has moved the cursor over the Imperfective Past tense form *prinimal* ‘accepted’ (see example (1) above) and the three options are given with iconic coloring: “Excellent” in dark green, “Acceptable” in light green, and “Impossible” in red. The participant needs only to click on one of the options and then move on. The participant is asked to rate both the Perfective and the Imperfective verb forms in each test pair.

There was no time limit on the experiment and participants were allowed to go back to items as many times as they wanted. They were prompted to go back and finish any items that they had skipped before they submitted their survey. Participants were also allowed, but not required, to write comments about their experience with the survey. After they submitted a complete survey, they were given a lottery code and the URL of a website where one half of each winning code would be displayed. They could then send the other half of their winning code to claim their prize.

**2.3 Recruitment of Participants**

Participants were recruited via emails sent to various listservs and individuals in Russia. Potential participants were told that they would have a chance to win a gift certificate worth 5000 rubles redeemable at a major Russian online store, and that there would be three winners for every 100 participants, with a maximum of 500 participants. They were given a brief description of the survey and told that only native speakers of Russian 16 and older would be allowed to participate and that no IP addresses or personal information would be collected. They were told that the purpose of the experiment was strictly scientific, that the goal was to learn more about Russian aspect in order to improve linguistic description and language pedagogy. We also provided our names, brief descriptions of who we are, and our websites.

In the course of one week (13.-20.09.2016), 501 participants successfully completed the task. As shown in Table 2, the number of responses per stimulus text ranges from 72 for the Scientific-Technical prose text ‘Bacteriophages’ (which is both relatively long and more difficult to read) to 99 for the Spoken Narrative text ‘Yellow Sign’. Altogether we collected over 55,000 data points.

The three options were converted to numerical scores of “Impossible” = 0, “Acceptable” = 1, and “Excellent” = 2.[[8]](#footnote-8)

Because there was some concern that participants might just click at random in order to get their lottery codes (despite the warning in the instructions), we sought to eliminate any outliers from the analysis. This was done by means of the following procedure. First the average response for each test item was computed across all participants. Then, for each participant, the deviation of each response from the group mean was computed and squared (so that positive and negative deviations would not cancel each other out). Next the average of the squared deviations of each participant was calculated and then the square root of that number was taken. Conceptually, this figure represents each participant’s average deviation from the group mean in their evaluation of the test items. Finally, the distribution of average deviations of the participants for each stimulus text were evaluated according to the standard definition of a statistical outlier as a value that exceeds 1.5 times the interquartile range of the distribution.[[9]](#footnote-9) In this way from one to four outliers were identified among the participants for each stimulus text, as shown in Table 2, yielding a total of 15 outliers, or less than 3% of the total population of participants. These outliers are excluded from further analysis. Table 5 displays the distributions of age and gender among the remaining participants.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
|  | Age distribution | | | | Gender counts | | |
|  | min | median | mean | max | Female | Male | Other |
| Beetle | 16 | 26.0 | 31.16 | 78 | 65 | 17 | 0 |
| Summit | 17 | 26.5 | 29.69 | 62 | 64 | 16 | 0 |
| Phages | 17 | 26 | 29.10 | 67 | 45 | 24 | 2 |
| Yellow Sign | 16 | 23 | 27.37 | 67 | 71 | 22 | 2 |
| MSLU | 17 | 27.5 | 29.11 | 64 | 48 | 27 | 1 |
| Ivan D | 17 | 24.0 | 29.85 | 62 | 58 | 24 | 0 |
| ALL | 16 | 25.0 | 29.33 | 78 | 351 | 130 | 5 |
| Table 5: Age and gender of participants | | | | | | | |

**3. Results and Discussion**

We first check the ratings of our experiment participants against the corpus frequency of the Perfective vs. Imperfective verbs in order to discover whether the ratings are influenced by frequency. We show in 3.1 that frequency effects, if any, are minimal. In 3.2 we explore our data in terms of what they tell us about the relationship between construal and redundancy. We find a gradual scale running from the majority of contexts where redundancy is high and construal has low saliency to those where construal is relatively free and redundancy is low. Section 3.3 focuses on the test items in the environment of known trigger words (see Table 2 in Section 1.3) in the texts, and shows that test items with triggers do not behave differently from the test items without triggers. In Section 3.4 we find that participants are more consistent in their rating of the verb form representing the original aspect in the text than in rating the verb form of the opposite aspect.

**3.1 Why frequency is not a factor**

An important first step in the analysis is to assess the influence of frequency on the rating of Perfective vs. Imperfective verb forms. If it is the case that native speakers have a strong tendency to simply choose the most frequent of the two verb forms, then our results would not be very interesting. As we show, however, the effect of frequency is minimal and this effect actually disappears when one removes from the data the inherent frequency biases for Perfective and Imperfective verb forms.

In Figure 2, the ratings are averaged over all participants (each item was rated by between 71 and 95 participants, as shown in the rightmost column in Table 2), and those weighted averages ranged from 0 (“Impossible”) to 2 (“Excellent”). The weighted average ratings are represented on the x-axis. The y-axis of Figure 2 shows the relative frequency of the tokens. Relative frequency is measured by taking the frequency (in the Russian National Corpus) of the token being rated divided by the frequency of the alternative token (the verb form with the opposite aspect) and then logarithmically transforming that ratio to create a “logit”.[[10]](#footnote-10) This calculation can be illustrated by the first verb pair that we encounter in example (1)(b): *prinjal* / *prinimal* ‘accepted’. The first item to be rated is the Perfective verb form *prinjal* (which also happens to be the original token), whereas the verb form of the opposite aspect is the Imperfective verb form *prinimal*. The relative frequency is thus the frequency of the token *prinjal* divided by the frequency of the alternative token *prinimal*. Since the frequency of *prinjal* is 73.81 per million and the frequency of *prinimal* is 28.19 per million, then the relative frequency is 2.62. Taking the log of 2.62 gives us the logit 0.96. In this case, the logit is a positive number because *prinjal* is more frequent than *prinimal*. For the rating of *prinimal*, the frequency relationship is the reverse, and because *prinimal* is less frequent than *prinjal*, the logit is a negative number, -0.96. Thus, positive logits on the y-axis indicate that the item being rated is more frequent than the alternative, whereas negative logits indicate that the alternative is more frequent than the item being rated. This arrangement of x- and y-axes means that the quadrants of Figure 2 show us the combinations of values described in Table 6.

|  |  |
| --- | --- |
| **Top Left**  Items that receive **low rating** and  are of **higher frequency** than alternative | **Top Right**  Items that receive **high rating** and  are of **higher frequency** than alternative |
| **Bottom Left**  Items that receive **low rating** and  are of **lower frequency** than alternative | **Bottom Right**  Items that receive **high rating** and  are of **lower frequency** than alternative |

Table 6: Combinations of values across the quadrants of Figure 2

Red dots in Figure 2 show ratings of original verb forms (the aspect that was actually used in the original text), while blue dots show the non-original tokens (verb forms of the opposite aspect). For some items, the dots have been replaced by labels of the same color. The “C” label stands for “control” and marks items where there is an absolute grammatical constraint that allows only one aspect to be expressed (control items appear on the extreme left and right edges of Figure 2). The “Tp” and “Ti” labels mark items where there are trigger words present that normally present cues for Perfective or Imperfective aspect (see Table 1). See Section 3.3 for more discussion of the distribution and behavior of control and trigger items.

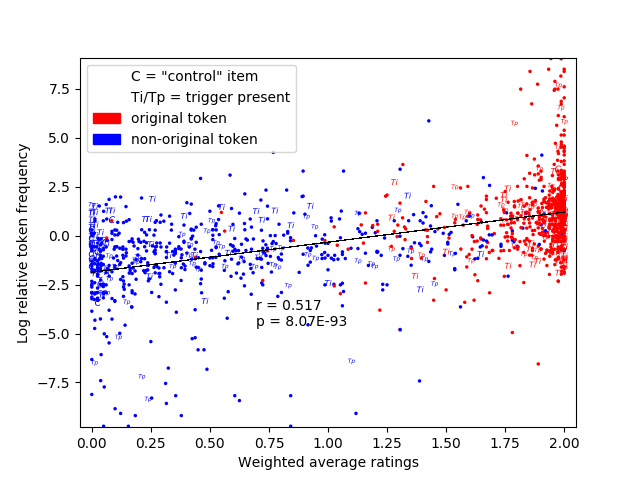


Figure 2: Weighted average ratings plotted against corpus frequencies of verb forms

Figure 2 shows that the native speakers mostly give the original token a higher rating, correctly identifying the original aspect based only on context, since most of the red dots are to the right (with higher average ratings) and most of the blue dots are to the left (with lower average ratings), although the separation of red originals from blue non-originals is far from perfect. Figure 2 shows a moderate and significant positive correlation between ratings and relative frequencies (r = 0.517, p = 8.07e-93). This seems to suggest that participants tend to select the token that is of higher frequency overall. But clearly frequency is not the whole story: if it was, r would be at or near 1 and all the points in the scatterplot would be hugging a line at a 45-degree angle. Instead, we see that the overwhelming mass of the data is very near 0 on the y-axis, where there is little or no difference in frequency between the two alternatives, so there is more to explain here.

It is possible that the apparent correlation in Figure 2 is distorted by another factor, namely the fact that Perfective and Imperfective verb forms are distributed very differently in general in Russian. Janda & Lyashevskaya (2011) showed, on the basis of nearly six million verb forms extracted from the Russian National Corpus, that Perfective Past tense forms are more than twice as frequent (1,972,287) as Imperfective Past tense forms (915,374). The difference in distribution of Perfective vs. Imperfective verb forms is significant (chi-square = 947756,

df = 3. p < 2.2e-16) with a medium-large effect size (Cramer’s V = 0.399). This distribution is visualized in Figure 3.

Figure 3: Distribution of verb forms across subparadigms, based on data from Janda & Lyashevskaya 2011[[11]](#footnote-11)

For example, when speakers give higher ratings to Perfective Past forms, it might seem that speakers are showing a preference for the more frequent forms, but the reality is that Perfective Past forms are on the whole more frequent than Imperfective Past forms to begin with. In other words, speakers will have more chances to rate Perfective Past forms that were the original forms in the texts than to rate Imperfective Past forms that were the original forms in the texts. In fact, given the distribution of original forms in the texts shown in Table 4, the participants get more than twice as many chances to rate Perfective Past forms that match the original (298) than Imperfective Past forms that match the original (134). Since there is a bias toward giving higher ratings to the forms that were in the original and since the Perfective Past is also more frequent than the Imperfective Past, the fact that speakers rated more Perfective Past matching forms than Imperfective Past matching forms compounds the apparent effect of frequency on rating. This magnification of the frequency effect occurs across nearly all the combinations of aspect and tense/mood in our experiment. Only in the Imperative are the two aspects represented in equal numbers (see Table 4), but there are only 20 test pairs in total for the Imperative, which thus comprises less than 3% of the test pairs. For all other tense + aspect/mood combinations, the higher frequency type of forms, namely the Perfective Past, Perfective Future[[12]](#footnote-12), and Perfective Infinitive are also the forms matching the original texts that were rated the most times, magnifying any frequency effect across the board.

We are able to partially control for this overall skew in the distribution of Russian verb forms by measuring the frequency at the lemma level (adding together the frequencies of all forms of each verb) instead of at the subparadigm level (as in Figure 2). In other words, we can look instead at the frequency of the whole verb, calculating the relative frequency as the total frequency of the lemma of the form being rated over the total frequency of the lemma of the corresponding form of the opposite aspect. Going back to example (1)(b): *prinjal* / *prinimal* ‘accepted’, the total frequency of the lemma *prinjat’* ‘accept’ is 217.37 per million. This number divided by the total frequency of the alternative lemma *prinimat’* ‘accept’ 157.99 per million, yields 1.38, and the log of that number is the logit 0.32. For *prinimal*, the frequency relationship is the reverse, and the logit is -0.32. This does not erase the effect of overall skew entirely, since there were over twice as many original Perfective verb forms in our test materials than original Impefective forms, but using lemma level frequency does mitigate the differences in distribution across the two aspects. As shown in Figure 4, the frequency effect practically disappears (r = 0.067) if we look at the total frequencies of verbs rather than specific forms, indicating that frequency is not a main effect, but interacts with inflectional morphology (as expected based on Janda & Lyashevskaya 2011).

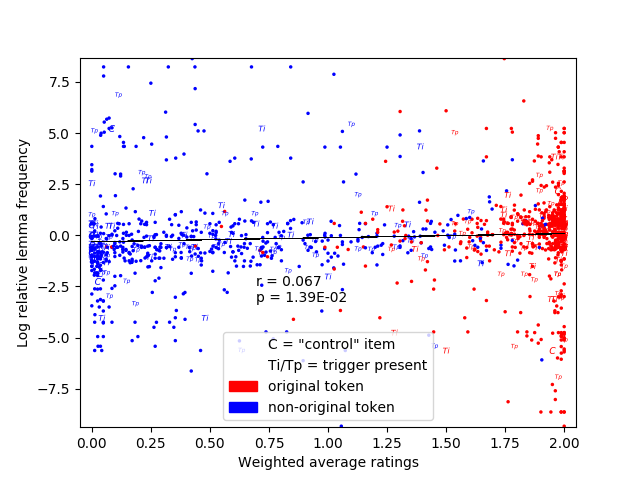


Figure 4: Weighted average ratings plotted against corpus frequencies of verb lemmas

**3.2 Redundancy vs. Construal**

Figure 5, like Figures 2 and 4, compares the average rating of the original token (the aspect that was actually used in the original text) to the average rating of the non-original token (the aspect opposite to the one in the original text), but in this figure, each dot represents both ratings. 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 dots represent original Imperfectives. Again, the labels “C”, Tp, and Ti replace dots in cases where there was a control item (no variation allowed), a trigger for Perfective, or a trigger for Imperfective.

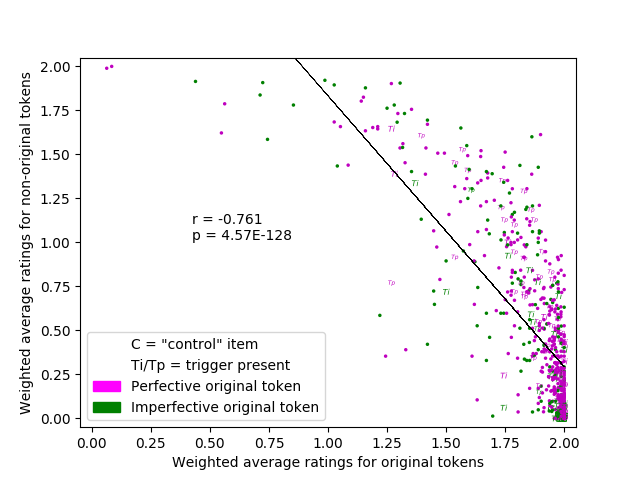


Figure 5: Comparison of weighted average ratings for original vs. non-original tokens

Table 7 indicates the combinations of parameters that appear in Figure 5, as well as the portions of data represented in each quadrant.

|  |  |
| --- | --- |
| **Top Left**  **Original** receives **low** rating  **Non-Original** receives **high** rating  Participants disagree with authors  2% of data | **Top Right**  **Original** receives **high** rating  **Non-Original** receives **high** rating  **CONSTRUAL**  17% of data |
| **Bottom Left**  **Original** receives **low** rating  **Non-Original** receives **low** rating  (No data) | **Bottom Right**  **Original** receives **high** rating  **Non-Original** receives **low** rating  **REDUNDANCY**  81% of data |

Table 7: Combinations of values across the quadrants of Figure 5

There is a strong and significant negative correlation (r = -0.76, p = 4.57e-128) between the two ratings for each item, which is what we would expect since it is reasonable that if a participant gives a high rating for one option, they will likely give the alternative a lower rating. However, the data is not evenly distributed, so the correlation alone does not tell us the whole story.

Most of the data (81%) lands in the bottom right quadrant, where the original token receives a high average rating, while the corresponding item of the opposite aspect receives a low average rating. At the extreme bottom right corner of the graph (where the original gets an average rating of 2 and the non-original gets an average rating of 0), aspect is fully redundant, since native speakers can reliably recover it based only on context. And aspect is relatively redundant for the remainder of the points in the bottom right quadrant as well. The construal of the event as Perfective or Imperfective is certainly present, but the overt marking of aspect on the verb is unnecessary since aspect can be recovered from context. Examples (2) and (3) from our experiment illustrate situations in which the Perfective and Imperfective aspect are highly redundant, with the original verb form marked with an asterisk. Table 8 shows the ratings that these test items received in our experiment. The ratings for the original aspect are in shaded boxes.

(2) *V vosem’ let mal’čik [ originalsbeža-l / non-originalsbega-l ] iz dom-a.*

[in eight.Acc year.Gen.Pl boy.Nom.Sg [ run.Perf-Past.Masc.Sg / run.Imperf-Past.Masc.Sg] from home-Gen.Sg]

‘At the age of eight the boy **ran away** from home.’

(3) *Bogomol’n-aja ženščin-a nikogda ne [ non-originalobruga-l-a / originalruga-l-a ] ego*

[pious-Nom.Sg.Fem woman-Nom.Sg never not [ yell.Perf-Past-Fem.Sg / yell.Imperf-Past-Fem.Sg ] he.Acc

‘The pious woman never **yelled** at him’

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Impossible = 0 | Acceptable = 1 | Excellent = 2 | Average |
| (2) Perfective (original aspect) | 0 | 0 | 82 | 2.0 |
| (2) Imperfective (non-original aspect) | 68 | 12 | 2 | 0.195 |
| (3) Perfective (non-original aspect) | 79 | 3 | 0 | 0.037 |
| (3) Imperfective (original aspect) | 0 | 0 | 82 | 2.0 |

Table 8: Ratings of examples (2) and (3) where aspect is redundant

Example (2) describes a unique punctual event that happened at a specific moment in time, which makes Perfective aspect strongly preferred. The categorical negation (literally ‘never not’) in (3) gives a strong reason to prefer the Imperfective for a situation that is designed to cover all times without limits (rather than specifying a unique event).

Most of the rest of the data (17%) is in the top right quadrant, where both the original and the non-original tokens receive high average ratings. While there are no points in the extreme top right (which would indicate that both choices are excellent), there are some points in this quadrant where both original and non-original tokens received very similar ratings (approaching a 45-degree line from the origin to the top right). For these items, construal is crucial, since native speakers cannot guess the aspect from context. Examples (4) and (5) illustrate data points from the top right quadrant where construal is more or less open, as evidenced in the rating data in Table 9.

(4) *On ume-l nezametno [ originalvytašči-t’ / non-originalvytaskiva-t’ ] den’gi iz karman-a zevak-a.*

[he.Nom know.how.Imperf-Past.Masc.Sg unnoticed [ pluck.Perf-Inf / pluck.Imperf-Inf] money.Acc from pocket-Gen.Sg idler-Gen.Sg]

‘He knew how **to** **pluck** the money out of the pocket of an idle onlooker without being noticed.’

(5) *Vyži-vš-uju iz um-a starux-u nikto vser’ez ne [ non-originalprinja-l / originalprinima-l ].*

[outlive-PastActivePart-Acc.Sg.Fem from mind-Gen.Sg old.woman-Acc.Sg no.one.Nom seriously not [ accept.Perf-Past.Masc.Sg / accept.Imperf-Past.Masc.Sg ]

‘No one **took** the senile old woman seriously.’

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Impossible = 0 | Acceptable = 1 | Excellent = 2 | Average |
| (4) Perfective (original aspect) | 10 | 40 | 32 | 1.268 |
| (4) Imperfective (non-original aspect) | 2 | 25 | 55 | 1.646 |
| (5) Perfective (non-original aspect) | 9 | 30 | 43 | 1.415 |
| (5) Imperfective (original aspect) | 4 | 25 | 53 | 1.598 |

Table 9: Ratings of examples (4) and (5) where aspect is open to construal

The use of aspect in example (4) depends on whether the speaker wishes to emphasize how successful the pickpocket was at snatching money (Perfective) or how continuously he stole money (Imperfective). In example (5) the Perfective construal emphasizes a single event (in this case, ignoring the woman when she accused her son-in-law of stealing from her), whereas the Imperfective emphasizes more a general disregard for the claims of a senile old woman. All of the ratings are very similar, in the range of about 1.3-1.6. In example (4), participants actually gave the Imperfective (which does not match the original in the text) a higher average rating than the (original) Perfective, and the difference in distribution of ratings is both significant and robust, but in example (5) there is no statistically significant difference in the distribution of ratings.[[13]](#footnote-13)

The remainder of the data (almost 2%) is in the top left quadrant, where it seems that most native speakers simply disagree with the authors of the texts, assigning the opposite values to the two options. Examples (6) and (7) illustrate test items from the top left quadrant, and their ratings are summarized in Table 10.

(6) *I govorit “Ladno, pošli!” Nu i my pošli. [ originalPošli / non-originalŠli ] my [ originalpošli / non-originalšli ], ja už tak v principe ponjala, čto estestvenno do konca my peškom ne dojdëm.*

[and say.Imperf-Pres.3.Sg alright go.Perf-Past.Pl. So and we go.Perf-Past.Pl. [ go.Perf-Past.Pl / go.Imperf-Past.Pl ] we [ go.Perf-Past.Pl / go.Imperf-Past.Pl ], I.Nom already thus in principle understand.Perf-Past.Fem.Sg that naturally to end we.Nom by-foot not go-all-the-way.Perf-Fut.1.pl. ]

‘And says “alright, let’s go!” So off we went. We set off, walked a little, I already knew that we wouldn’t make it all the way very easily by foot.’

(7) *Snačala ja zapolnil anketu, prišël tuda, xotel uže [ non-originalsdat’ / originalsdavat’ ], no mne skazali, čto tam čto-to, čego-to ne xvataet, čto, nu, kak vsegda.*

[first I fill-out-Perf-Past.Masc.Sg form, come-Perf-Past.Masc.Sg there, want.Imperf-Past.Masc.Sg already [ submit.Perf-Inf / submit.Imperf-Inf ], but I.Dat tell.Perf-Past.Pl, that there something.Nom, something.Gen not sufficient.Imperf-Pres.3.Sg, which, well, as always]

‘First I filled out the form, walked up and wanted to just **submit** it, but I was told that this and that was missing, which, you know, is how it always is.’

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Impossible = 0 | Acceptable = 1 | Excellent = 2 | Average |
| (6) Perfective (original aspect) | 88 | 6 | 1 | 0.084 |
| (6) Imperfective (non-original aspect) | 0 | 0 | 95 | 2.000 |
| (7) Perfective (non-original aspect) | 0 | 6 | 70 | 1.921 |
| (7) Imperfective (original aspect) | 19 | 39 | 18 | 0.987 |

Table 10: Ratings of examples (6) and (7) where participants disagree with authors

There are no test items for which native speakers found neither form (original, non-original) to be suitable in the given context, and thus no points in the bottom left quadrant.

The distribution in Figure 5 is stable across the four subparadigms (see Table 4), as shown in Table 11. It is also stable across genres, as shown in Table 12.

|  |  |  |
| --- | --- | --- |
|  | r (effect size) | p (statistical significance) |
| **Past** | -0.773 | 3.88e-87 |
| **Future** | -0.709 | 6.69e-10 |
| **Infinitive** | -0.783 | 3.70e-35 |
| **Imperative** | -0.661 | 1.51e-3 |
| ***Overall (as in Figure 5)*** | -0.761 | 4.57e-128 |

Table 11: Comparison of correlations across subparadigms

|  |  |  |
| --- | --- | --- |
|  | r (effect size) | p (statistical significance) |
| **Beetle** | -0.860 | 5.20e-45 |
| **Summit** | -0.754 | 2.01e-16 |
| **Phages** | -0.837 | 4.30e-27 |
| **Yellow Sign** | -0.765 | 1.52e-16 |
| **MSLU** | -0.795 | 1.67e-31 |
| **Ivan D** | -0.659 | 1.66e-16 |
| ***Overall (as in Figure 5)*** | -0.761 | 4.57e-128 |

Table 12: Comparison of correlations across genres

Figure 5 tells an important story about redundancy and construal. The meanings available for construal have to come from somewhere, so it is reasonable that they are anchored in uses where the construal, though of course present, is redundant. In a usage-based model, this means that the construal associated with, say, the Perfective aspect, is entrenched through repeated exposure to examples where the construal is unmistakable due to the presence of other clues in the context that align with that construal. This entrenchment makes the construal of Perfective robust enough to be capable of expressing Perfective meaning even when there are no other clues to Perfective available in the context. However, these two characterizations, one of a situation where aspect markers are entirely redundant, and the other of free construal of aspect so that the speaker can choose what to emphasize, are extreme idealizations. In reality, there are no distinct types of situations in this data, no clear groupings at all. While there are certainly many examples where one aspect is strongly preferred (making aspect marking entirely or almost entirely redundant), and some examples where both aspects are acceptable (meaning that aspect is largely a matter of construal), there are no clear divisions between these types. Instead, nearly all examples involve some balance between redundancy and construal, with most examples reflecting a heavier portion of redundancy than construal. Redundancy is the norm, and probably sets the standard for what the usually redundant aspect markers can express on the occasions when they have the opportunity to appear on their own. While it remains to be proven, it is conceivable that this kind of relationship between redundancy and construal is common among languages of the world.

**3.3 Controls and Triggers**

We designated 14 tokens in our texts as “control” items because their grammatical contexts disallow one aspect or another. For example, the Imperfective aspect is required for infinitive complements of phasal verbs, as in *načinaet* ***rasti*** ‘begins **to grow**’.

We also labeled items in which the verb was collocated with one of the trigger words that we collected from Russian textbooks.[[14]](#footnote-14) Table 13 summarizes the number of items in each category.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Total number of test item pairs | Controls | Triggers associated with Perfective | Triggers associated with Imperfective | Triggers (total) |
| **Beetle** | 150 | 5 | 2 | 22 | 24 |
| **Summit** | 83 | 1 | 10 | 2 | 12 |
| **Phages** | 99 | 1 | 19 | 14 | 33 |
| **Yellow Sign** | 80 | 2 | 6 | 4 | 10 |
| **MSLU** | 139 | 2 | 10 | 6 | 16 |
| **Ivan D** | 122 | 3 | 8 | 9 | 17 |
| ***Overall (as in Figure 5)*** | 673 | 14 | 55 | 57 | 112 |

Table 13: Distribution of controls and triggers across genres

Table 14 shows the mean ratings assigned by participants to control items, items associated with triggers, and all remaining items.

|  |  |  |  |
| --- | --- | --- | --- |
|  | Number of test item pairs | Mean rating  (original aspect) | Mean rating  (non-original aspect) |
| Controls | 14 | 1.9830 | 0.0198 |
| Triggers associated with Perfective | 55 | 1.8658 | 0.5985 |
| Triggers associated with Imperfective | 57 | 1.9158 | 0.2570 |
| Triggers (total) | 112 | 1.8912 | 0.4247 |
| Remaining test items (excluding controls and triggers) | 561 | 1.8414 | 0.5453 |

Table 14: Ratings of “control” items and items associated with trigger words

As expected, the control items exhibit near-categorical ratings, with original tokens receiving mean ratings only slightly less than the maximum score of 2.0, and non-original tokens receiving mean ratings only slightly over the minimum score of 0.0.

For items with triggers indicating an overt preference for one aspect over the other, it would be natural to expect ratings that are more categorical than items without triggers, but Table 14 shows that the ratings of items with triggers are very comparable to the ratings of items without triggers.

A scatterplot of ratings for items with triggers is given in Figure 6, which has the same x- and y-axes as Figure 5, but shows only the data for the items with triggers. Table 15 shows the distribution of trigger test items in terms of the numbers of items that appear in the top right and bottom right quadrants of Figure 6. On Figure 6, a purple “p” shows the average rating of a test item pair where the trigger should predict a Perfective verb and there was a Perfective verb in the original text. A green “p” shows the average rating when the trigger should predict a Perfective, but there was an Imperfective in the original text. These two combinations are represented in the top row of Table 15, with the combination where the trigger prediction and the original text are the same (both Perfective) in the shaded boxes. A purple “i” on Figure 6 represents a test pair where the trigger would predict an Imperfective, but the original text has a Perfective. And a green “i” shows items for which both the trigger prediction and the original text are Imperfective. These two combinations are shown in the bottom row of Table 15, again with the items where the prediction and the original aspect match in shaded boxes. Table 15 further shows the breakdown of trigger test pairs according to whether the mean rating for the non-original was above 1.0 (in the top right quadrant of Figure 6, where construal is relatively free) or below 1.0 (in the bottom right quadrant of Figure 6, where aspect is mostly redundant).

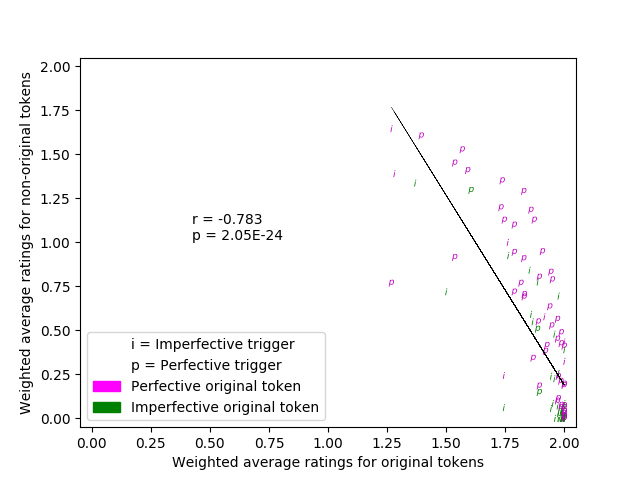


Figure 6: Weighted average ratings for original vs. non-original tokens in trigger items

|  |  |  |
| --- | --- | --- |
|  | Original token is **Perfective** | Original token is **Imperfective** |
| Triggers associated with **Perfective** | **11**  purple “p” in top right | **1**  green “p” in top right |
| **41**  purple “p” in bottom right | **2**  green “p” in bottom right |
| Triggers associated with **Imperfective** | **2**  purple “i” in top right | **1**  green “i” in top right |
| **6**  purple “i” in bottom right | **48**  green “i” in bottom right |

Table 15: Distribution of test item pairs with triggers on Figure 6

The distribution of items collocated with triggers is overall very similar to that of items without triggers: 13% of items with triggers land in the top right quadrant, while 87% land in the bottom right quadrant. This performance is only marginally better than that of the test items overall (compare Table 7), suggesting that there are probably many other cues that native speakers access from context, but have not been adequately described or cannot be reduced to simple collocations. The identification of these contextual cues goes beyond the scope of the present article (but see suggestions for future research in Section 4).

**3.4 Agreement/Disagreement Across Speakers**

There is mounting evidence that, instead of converging on a single grammar, native speakers can disagree on what is grammatical in their language and differ widely in their attainment of their native language. The present study contributes to this evidence, and further shows that differences in acceptability ratings are much more pronounced when speakers are presented with unattested language (items that do not match authentic original texts) than with attested language.

Dąbrowska (Dąbrowska 2008, 2012, 2013, 2015; Street & Dąbrowska 2010) has shown, through a series of experiments on native speakers of Polish and English, that native speakers exhibit differences in their grammars. Native speakers can have different strategies for understanding the same grammatical phenomena (for example abstract schematic rules vs. low-level rules vs. memorization of exemplars), yet still produce the same forms, and thus be said to “speak the same language”. Native speakers can also differ in how well they master the grammatical categories of their native language. These differences span various kinds of grammatical phenomena, including morphology and syntax. Dąbrowska (2015: 661-662) attributes such individual differences to both cognitive and environmental factors. In other words, differences can result from differences in inherent ability and also from differences in the language that people are exposed to: the precise input of course varies from person to person, and there are overall differences in the quantity and quality of language exposure that are tied to socio-economic status.

The data described in sections 3.1-3.3 shows that native speakers of Russian can vary widely in their rating of the acceptability of Perfective vs. Imperfective aspect in context. In our data, the extent of such variation is partly connected to the status of the item being rated. When native speakers rated the item that originally appeared in the text (the aspect matching that in the authentic text), they tended to agree that the item was “Excellent”. We see this, for example, in the ratings in Table 8 for examples (2) and (3). In example (2), the original aspect was Perfective, and all 82 participants rated the Perfective form as “Excellent”. In example (3), the original aspect was Imperfective, and all 82 participants rated the Imperfective form as “Excellent”. However, when native speakers rated the non-original aspect (the aspect opposite to the one in the authentic text), they often chose a wide range of ratings, as we see in the ratings of non-matching items in Tables 8 and 9. Example (8) and Table 16 show an extreme example of lack of agreement among participants for the non-original verb form.

(8) *Fag-ov [ podverg-l-i / \*podverga-l-i ] polnogenomn-omu sekveknirovani-ju*

[phage-Acc.Pl [ subject.to.Perf-Past-Pl / subject.to.Imperf-Past-Pl ] full.gene-Dat.Sg.Neut sequencing-Dat.Sg]

‘The phages **were subjected to** full-gene sequencing’

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Impossible = 0 | Acceptable = 1 | Excellent = 2 | Average |
| (8) Perfective (non-original aspect) | 24 | 24 | 23 | 0.986 |
| (8) Imperfective (original aspect) | 2 | 13 | 56 | 1.761 |

Table 16: Ratings of example (8)

While speakers mostly agreed that the original aspect in example (8), the Imperfective, was “Excellent” or “Acceptable”, the rating of the non-original aspect, the Perfective, is completely split across the three options. And the difference in rating distributions for the Perfective vs. Imperfective is both significant and large.[[15]](#footnote-15)

Figure 7 visualizes the difference in distribution of average ratings between the original items and the non-original items. The majority of average ratings for the original items are at or near 2.0, which is a perfect score, giving a very skewed distribution, whereas the distribution of the average ratings for the non-original items is much wider.

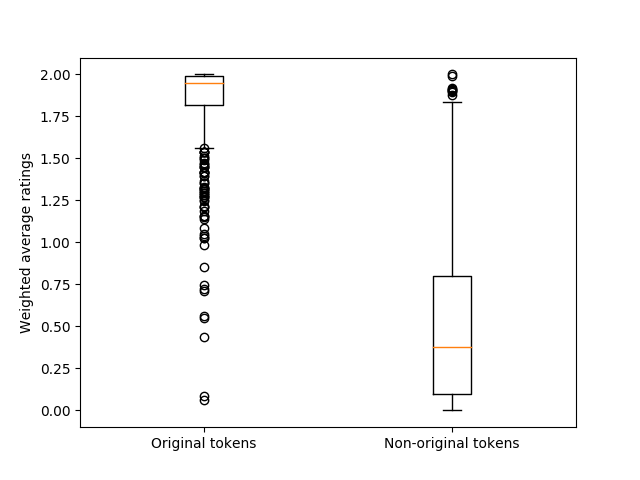


Figure 7: Distributions of average ratings for original vs. non-original items

One way of measuring the degree of agreement/disagreement across participants is to look at the standard deviation in the rating of each item, since a larger standard deviation will indicate greater diversity of responses. Figure 8 visualizes the distributions of standard deviations for the responses to the original tokens vs. the non-original tokens of the opposite aspect, showing that responses to the originals were different from those to the non-originals also on this measure. The mean standard deviation for original items was 0.28, whereas the mean standard deviation for the non-original items was 0.49, and a paired t-test shows that this difference is significant (test-statistic = -17.26, p = 1.5e-55).

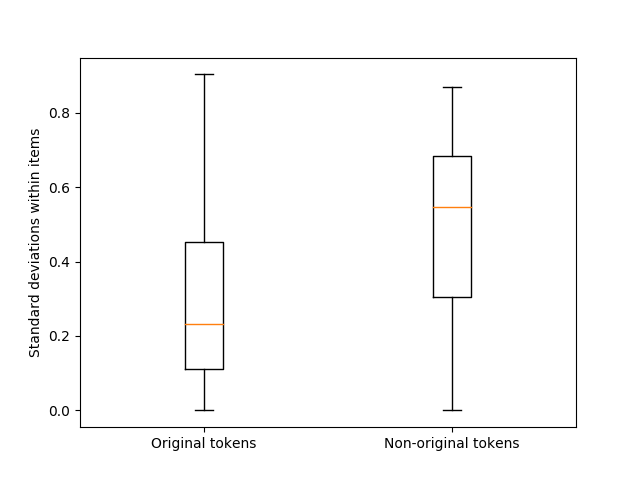


Figure 8: Distributions in standard deviations of ratings for original vs. non-original items

In addition, as shown in the top left quadrant of Figure 5, for about 2% of our test item pairs, the participants in our experiment tend to disagree with the authors of the texts about which aspect is more suitable. Such examples give further evidence of divergence in the grammars of speakers.

We find that native speakers can differ in their ratings of Perfective and Imperfective verb forms in Russian. This finding is in line with other experimental evidence showing individual variation in the grammars of native speakers. In addition, we find that native speakers are more consistent in giving positive ratings for the original tokens, whereas they are less consistent in their ratings of the non-original tokens (which get larger standard deviation and variance). This may indicate that native speakers are more reliable in reacting to authentic language, than in reacting to language that has been manipulated (in this case, by suggesting an aspectual form that does not match the original text). This result may also have implications for how much linguists can rely on the intuitions of native speakers in reaction to constructed “examples” as opposed to authentic ones.

**4. Conclusions**

Aspect is one of the most pervasive and characteristic grammatical categories in Russian, requiring its speakers constantly to choose between Perfective and Imperfective. Our experiment on the rating of Perfective and Imperfective verb pairs in full, authentic contexts, shows that this choice is anything but simple and invariable. We find that in about 81% of examples, native speakers can fairly reliably retrieve the original aspect, and that they can do this regardless of whether there is an identifiable “trigger” word for aspect in the context. In about 17% of contexts, native speakers accept both aspects. While there are no distinct groups in this data, the discovery of these distributions and the norming of concrete examples along the scale from categorical grammaticality difference (correct/incorrect) to free variation is a valuable contribution to our knowledge about the behavior of aspect in Russian. These findings can serve as the basis for further research using methods of experimentation and machine learning to ferret out the as-yet unidentified contextual clues to aspect. If such clues can be uncovered, this could have far-reaching implications for both natural language processing and language pedagogy.

The data from our experiment reveals in a concrete way the relationship between redundancy and construal, which has not previously been studied empirically. The distribution of ratings in Figure 5 is compatible with the interpretation of redundancy and construal as co-existing in a continuum. The meanings associated with the two alternative construals offered by Russian aspect are always available, but their salience and independence from context vary. At one end of the continuum, the meanings of Perfective vs. Imperfective are strongly anchored by context, highly redundant, and the choice of aspect is tightly constrained. This end of the continuum is also its center of gravity, the place where most uses are observed and most entrenchment is expected. At the other end of the continuum, construal breaks free from context and operates independently, without the support of redundancy. In these uses, the speaker can deploy aspect to manipulate nuances of meaning, with the option of representing the “same” content in two slightly different ways, emphasizing either the discreteness of the situation as Perfective or its fluidity as Imperfective. Between these two extremes there are varying degrees of redundancy and freedom of construal. It is likely that construal actually needs the redundant uses to empower it to operate on its own when redundancy is reduced. Our data is restricted only to the relationship between redundancy and construal for Russian aspect; it remains to be seen whether this relationship can be confirmed for other grammatical distinctions and other languages.

Our data also confirms that there are differences among native speakers, and additionally reveals a consistent bias toward less variation in response to an attested (naturally occurring) example than to a non-attested (manipulated and possibly unnatural) example. This suggests that native speakers may be more reliable in reacting to authentic language stimuli than in reacting to language stimuli constructed for experimental purposes (even when such language is merely a slight modification of authentic language). This may mean that linguistic experiments involving constructed stimuli could be subject to a certain bias, but this is also a topic for future research.

**References**

Andrews, Edna, Galina N. Aver’janova, Galina I. Pjadusova. 2001. *The Russian Verb: Form & Function.* Moscow: Russkij jazyk.

Anketa. 1997. Anketa aspektologičeskogo seminara filologičeskogo fakul’teta MGU. Trudy aspektologičeskogo seminara filologičeskogo fakul’teta MGU, V. 2. Moscow: Izdatel’stvo MGU.

Avilova, Natal’ja S. 1968. Dvuvidovye glagoly s zaimstvovannoj osnovoj v russkom literaturnom jazyke novogo vremeni. *Voprosy jazykoznanija*: 66-78.

Bartoň, Tomáš, Václav Cvrček, František Čermák, Tomáš Jelínek, Vladimír Petkevič. 2009. *Statistiky češtiny*. Prague: Nakladatelství Lidové Noviny.

Batiukova, Olga, Pier Marco Bertinetto, Alessandro Lenci, Alessandra Zarcone. 2012. Semantic priming study of Russian aspect and resultativity. *Oslo Studies in Language* 4, 177-206.

Bresnan, Joan. 2007. Is syntactic knowledge probabilistic? Experiments with the English dative alternation. In: Sam Featherston, Wolfgang Sternefeld, eds. *Roots: Linguistics in Search of its Evidential Base*, 75–96. Berlin: Mouton de Gruyter.

Čertkova, Marina Jurevna. 1996. *Grammatičeskaja kategorija vida v sovremennom russkom jazyke*. Moscow: Moscow State University.

Chiari, Isabella. 2007. Redundancy Elimination: The Case of Artificial Languages. *Journal of Universal Language* 8, 7-38.

Croft, William and D. Alan Cruse. 2004. *Cognitive Linguistics*. Cambridge: Cambridge University Press.

Cubberly, Paul. 2002. Russian: A Linguistic Introduction. Cambridge: Cambridge University Press.

Čujkova, Oksana Ju. 2017. *Buduščee vremja v russkom jazyke: aspektual’nost’, temporal’nost’, modal’nost’ (korpusnoe issledovanie)*. MA Thesis, St. Petersburg State University. St. Petersburg, Russia.

Dąbrowska, Ewa. 2008. The later development of an early-emerging system: The curious case of the Polish genitive. *Linguistics* 46:3, 629–650.

Dąbrowska, Ewa. 2012. Different speakers, different grammars: Individual differences in native language attainment. *Linguistic Approaches to Bilingualism* 2: 219-253.

Dąbrowska, Ewa. 2013. Functional constraints, usage, and mental grammars: A study of speakers’ intuitions about questions with long-distance dependencies. *Cognitive Linguistics* 24(4): 633 – 665.

Dąbrowska, Ewa. 2015. Individual differences in grammatical knowledge. In: Ewa Dąbrowska, Dagmar Divjak, eds. *Handbook of Cognitive Linguistics*, 649–667. Berlin: De Gruyter Mouton.

Dahl, Östen. 2004. *The Growth and Maintenance of Linguistic Complexity.* Amsterdam: John Benjamins.

Dickey, Stephen M. 2000. *Parameters of Slavic Aspect: A Cognitive Approach.* Stanford, CA: CSLI Publications.

Dickey, Stephen M. and Laura A. Janda. 2015. Slavic Aspectual Prefixes and Numeral Classifiers: Two Kinds of Lexico-Grammatical Unitizers. *Lingua* 168, 57-84.

Endresen, Anna and Laura A. Janda. 2017. Five statistical models for Likert-type experimental data on acceptability judgments. *Journal of Design and Statistics in Linguistics and Communication Science*. (forthcoming)

Ellis, Nick C. and Stefanie Wulff. 2015. Second language acquisition. In: Ewa Dąbrowska, Dagmar Divjak, eds. *Handbook of Cognitive Linguistics*, 409-432. Berlin: De Gruyter Mouton.

Gagarina, N. 2004. Does the acquisition of aspect have anything to do with aspectual pairs? *ZAS Papers in Linguistics*, 33, 39-61.

Galton, Herbert. 1976. *The main functions of the Slavic verbal aspect*. Skopje: Macedonian Academy of Sciences and Arts.

Gladney, Frank Y. 1982. Biaspectual verbs and the syntax of aspect in Russian. *Slavic and East European Journal* (26): 202-215.

Gorbova, Elena V. 2010. Grammatičeskaja kategorija aspekta i kontekst (na materiale ispanskogo i russkogo jazykov). Doctoral Dissertation, St. Petersburg State University. St. Petersburg, Russia.

Isačenko, Aleksandr V. 1960. *Grammatičeskij stroj russkogo jazyka v sopostavlenii s slovackim*. *Morfologija, čast vtoraja.* Bratislava: Vydavatel’stvo Slovenskej akadémie vied.

Jakobson, Roman O. 1959/1971. Boas’ view of grammatical meaning. In: *Selected Writings*, Vol. 2 489-496. The Hague: Mouton.

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. 2007a. What makes Russian Bi-aspectual verbs Special. In: Dagmar Divjak and Agata Kochanska, eds. Cognitive Paths into the Slavic Domain*. Cognitive Linguistics Research*,83-109*.* Berlin/New York: Mouton de Gruyter.

Janda, Laura A. 2007b. Aspectual clusters of Russian verbs. *Studies in Language* 31:3, 607-648.

Janda, Laura A., 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., Anna Endresen, Julia Kuznetsova, Olga Lyashevskaya, Anastasia Makarova, Tore Nesset, Svetlana Sokolova. 2013. *Why Russian aspectual prefixes aren’t empty: prefixes as verb classifiers*. Bloomington, IN: Slavica Publishers.

Jászay, László. 1999. Vidovye korreljaty pri dvuvidovyx glagolax. *Studia Russica* (17): 169-177.

Langacker, Ronald W. 2008. *Cognitive Grammar*. Oxford: Oxford University Press.

Langacker, Ronald W. 2015. Construal. In: Ewa Dąbrowska, Dagmar Divjak, eds. *Handbook of Cognitive Linguistics*, 120-143. Berlin: De Gruyter Mouton.

Martelle, Wendy. 2011. *Testing the Aspect Hypothesis in L2 Russian*. Doctoral Dissertation, University of Pittsburgh.

Mučnik, I. P. 1966. Razvitie sistemy dvuvidovyx glagolov v sovremennom russkom jazyke. *Voprosy jazykoznanija*: 61-75.

Offord, Derek. 2005. *Using Russian: A guide to contemporary usage.* Cambridge: Cambridge University Press.

Padučeva, Elena V. 2008. Režim interpretacii kak kontekst, snimajuščij neodnoznačnost´. *Komp´juternaja lingvistika i intellektual´nye texnologii. Vyp. 7 (14). Po materialam meždunarodnoj konferencii “Dialog 2008”*, 412-419.

Polinsky, Maria. 2008. Without aspect. In: Greville G. Corbett, M. Noonan, eds. *Case and Grammatical Relations*, 263-282. Amsterdam: John Benjamins.

Reynolds, Robert J. 2016. *Russian natural language processing for computer-assisted language learning.* Doctoral Dissertation, UiT The Arctic University of Norway.

Stoll, Sabine. 2001. *The Acquisition of Russian Aspect*. Doctoral Dissertation, University of California, Berkeley.

Street, James and Ewa Dąbrowska. 2010. More individual differences in Language Attainment: How much do adult native speakers of English know about passives and quantifiers? *Lingua* 120, 2080-2094.

Talmy, Leonard. 1978. Figure and ground in complex sentences. In: Joseph Greenberg, ed. *Universals of Human Language* V. 4, 625-49. Stanford, CA: Stanford University Press.

Talmy, Leonard. 1996. Fictive motion in language and “ception”. In: P. Bloom, M. A. Peterson, L. Nadel, F. Garrett, eds. *Language and Space*, 211-276. Cambridge, MA: MIT Press.

Taylor, John. 2002. *Cognitive Grammar*. Oxford: Oxford University Press.

Timberlake, Alan. 2004. *A Reference Grammar of Russian*. Cambridge: Cambridge U Press.

Townsend, Charles E. 1975. *Russian Word-Formation*. Columbus, OH: Slavica Publishers.

Verhagen, Arie. 2007. Construal and Perspectivization. In: Dirk Geeraerts, Hubert Cuyckens, eds. *The Oxford Handbook of Cognitive Linguistics*, 48-81. Oxford: Oxford University Press.

Vinnitskaya, Inna and Ken Wexler. The role of pragmatics in the development of Russian aspect. *First Language* 21, 143-186.

Zaliznjak, Anna A., and Aleksej D. Šmelev. 2000. *Vvedenie v russkuju aspektologiju*. Moscow: Jazyki russkoj kul’tury.

Zipf, G. 1949. *Human Behavior and the Principle of Least Effort.* New York: Addison-Wesley.

**Appendix 1**

**List of reference grammars and textbooks from which “triggers” were assembled:**

Andrews, Edna, Galina N. Aver’janova, Galina I. Pjadusova. 2001. *The Russian Verb: Form & Function.* Moscow: Russkij jazyk.

Bogojavlensky, Marianna. 1982. *Russian Review Grammar.* Columbus, OH: Slavica.

Borras, Frank M. & Reginald F. Christian. 1959. *Russian Syntax: Aspects of Modern Russian Syntax and Vocabulary.* Oxford: Clarendon Press.

Brown, Nicholas J. 1996. *The New Penguin Russian Course.* London: Penguin Books.

Davis, Patricia Anne. 2009. Making Progress in Russian. Hoboken, NJ: Wiley.

Kagan, Olga, Frank Miller, & Ganna Kudyma. 2006. *V Puti. Russian Grammar in Context. 2nd ed.* Upper Saddle River, NJ: Prentice Hall.

Kagan, Olga, Tatiana Akishina, Richard Robin. 2002. *Russian for Russians.* Bloomington, IN: Slavica Publishers.

Lekic, Maria D., Dan E. Davidson, Kira Gor. 2008. *Russian Stage One: Live From Russia: Vol 1 & 2.* Dubuque, IA: Kendall Hunt.

Levine, James S. 1999. *Schaum’s outline of Russian Grammar*. New York: McGraw-Hill.

Lubensky, Sophia, Gerard L. Ervin, Larry McLellan, Donald K. Jarvis, et al. 2002. *Nachalo*. New York: McGraw Hill.

Murphy, Arthur B. 1965. *Aspectival Usage in Russian*. Oxford: Pergamon Press.

Nummikoski, Marita. 2012. Troika: a communicative approach to Russian language, life, and culture. Hoboken, NJ: Wiley.

Offord, Derek. 1993. *Modern Russian: An Advanced Grammar Course*. London: Bloomsbury.

Offord, Derek. 2005. *Using Russian: A guide to contemporary usage.* Cambridge: Cambridge University Press.

Pulkina, Il´za M., Ekaterina B. Zaxava-Nekrasova. 1978. *Učebnik russkogo jazyka, 6th ed.* Moscow: Russkij jazyk.

Rassudova, Ol´ga P. 1968. *Upotreblenie vidov glagola v russkom jazyke.* Moscow: Izdatel´stvo Moskovskogo universiteta.

Robin, Richard M., Karen Evans-Romaine & Galina Shatalina. 2013. *Golosa*. New York: Pearson Education.

Rosengrant, Sandra Freels. 2006. *Russian in Use.* New Haven, CN: Yale University Press.

Rifkin, Benjamin A. 1995. *Grammatika v Kontekste : Russian Grammar in Literary Contexts.* New York: McGraw Hill.

Timberlake, Alan. 2004. *A Reference Grammar of Russian.* Cambridge: Cambridge University Press.

Townsend, Charles E. 1970. *Continuing with Russian*. New York: McGraw-Hill.

Wade, Terence. 1992. *A Comprehensive Russian Grammar.* Oxford: Blackwell.

Zauber, Issa R. and Jan Zamir. 1997. *750 Russian Verbs and Their Uses.* New York: John Wiley & Sons.

**Appendix 2**

**Sources for texts used in the experiment:**

|  |  |  |
| --- | --- | --- |
| **Genre** | Title | Source |
| Fiction | *Besprizornik Žuk* / ‘Beetle, the Vagrant Boy’ | © 2015 Fineeva Elizaveta  Biblioteka Maksima Moškova |
| Journalistic Prose | *Počemu neftjanoj sammit v Doxe provalilsja* / ‘Why the Petroleum Summit in Doha Failed’ | Mixail Krutixin, Carnegie Moscow Center © 2016 |
| Scientific-Technical Prose | *Učenye vyjasnili, počemu bakteriofagam trudno borotsja s immunnoj sistemoj bakterij* / ‘Scientists have discovered why bacteriophages have a hard time battling with the immune system of bacteria’ | Aleksandr Markov, element.ru 18.04.2016 |
| Spoken Narration | *Istorija o tom, kak na našej obščažnoj kuxne pojavilsja tot samyj pervyj želten’kij znak* / ‘The story of how the first yellow sign got to the kitchen in our dorm’ | From the corpus: *Rasskazy o snovidenijax i drugie korpusa zvučaščej reči* / ‘Narrations of dreams and other oral corpora’ (А. А. Kibrik et al.) © 2016 |
| Guided Spoken Narration | Moscow State Linguistic University Video 3 | The Multimodal Communication and Cognition Laboratory at Moscow State Linguistic University (Alan Cienki, Olga Iriskhanova) © 2014 |
| Radio Interview | Ivan Dmitrievič | GTRK “Lipeck”. Broadcast from the *Vstreči* / ‘Meetings’ series, November 2004 |

1. See Chiari 2007 for a definition of linguistic redundancy in terms of information theory and scholarly overview. [↑](#footnote-ref-1)
2. Russian has several hundred “bi-aspectual” verbs that express both Perfective and Imperfective aspect (scholars disagree on exact numbers, cf. numerous citations in Janda 2007a). However, most scholars agree that in context, bi-aspectual verbs are not ambiguous: they always express either Perfective or Imperfective aspect (cf. Isačenko 1960: 143-44; Mučnik 1966: 61; Avilova 1968: 66; Galton 1976: 294; Gladney 1982: 202; Čertkova 1996: 100-109; Jászay 1999: 169; Zaliznjak and Šmelev 2000: 10; Janda 2007a). In other words, these verbs are syncretic aspectual pairs for which aspect is disambiguated in context, just as number is disambiguated for English *sheep* in context. The one dissenter from this interpretation is Timberlake (2004: 407-9) who suggests that bi-aspectual verbs are “anaspectual”, and do not express aspect, but he is in the minority. [↑](#footnote-ref-2)
3. Trigger words vary in their strength of preference for one aspect. With phasal verbs there is an absolute preference for Imperfectives, and therefore phasal verbs figure among the control items in our stimuli described in Sections 2-3. [↑](#footnote-ref-3)
4. Padučeva (2008) suggests that there are differences in the aspectual interpretation of Russian verbs across genres and genre-related differences in the aspectual behavior of verbs have been reported for Czech (Bartoň et al. 2009: 166-168), leading us to suspect that aspect might behave differently across genres in Russian. [↑](#footnote-ref-4)
5. There are approximately twenty (depending upon how one counts allomorphs) affixes that mark aspect in Russian, in addition to suppletion. A fuller description of Russian aspectual morphology can be found in Townsend 1975. [↑](#footnote-ref-5)
6. Janda & Lyashevskaya (2011) grouped their data according to the inflectional forms of verbs, which meant that Non-Past forms were labeled as such, rather than separating forms according to Present vs. Future tense, and the periphrastic Future was not represented as separate from the Imperfective Infinitive. However, they did find that the Perfective Non-Past (= Future) was less common than both the Past and Infinitive, but more common than the Imperative, again showing that the distributions in our stimuli conform to overall distributions in Russian. [↑](#footnote-ref-6)
7. <https://www.qualtrics.com/research-core/> [↑](#footnote-ref-7)
8. We are of course aware of the potential drawbacks of assigning numerical values to Likert-scale evaluations such as the one used in our experiment, particularly the fact that the distances between the evaluations on the scale might not be equal. However, there is growing evidence that the results of statistical tests where these values are treated as interval data yield very similar results to tests in which they are treated as categorical data (see Endresen & Janda 2017 for examples and extensive theoretical discussion of this issue). Furthermore, the large size of our experiment increases the likelihood that individual differences among participants would be corrected for by the sheer mass of data. [↑](#footnote-ref-8)
9. This was done using the R function boxplot.stats()$out. [↑](#footnote-ref-9)
10. It is customary to logarithmically transform corpus frequency data in order to correct for the extreme skewing of corpus word frequencies, known as “Zipf’s Law” (1949). Logits are logarithmically transformed odds ratios and are the basis for logistic regression. They have the admirable property of transforming odds ratios (which normally range from zero to 1 on one side, and from 1 to infinity on the other) into a symmetrical distribution. For example, an odds ratio of 1000/1 (relative frequency where one item appears 1000 times and the other only once) = 1000 yields a logit of 6.9, and the reverse relative frequency of 1/1000 = 0.001 yields a logit of -6.9. When the frequencies of two items are the same, the odds ratio is 1, and the logit is 0. [↑](#footnote-ref-10)
11. Note that the classification of verb forms in Janda & Lyashevskaya 2011 is based solely on verbal morphology, which was necessary to facilitate automated identification in such a large study. This means that the Perfective Future and Imperfective Present were both classified as Non-Past, since they have the same inflectional morphology. Also, the Imperfective Future is not separately identified in Janda & Lyashevskaya because it is a periphrastic tense, the two components of which can be separated my several words and appear in either order, making reliable automated identification virtually impossible. The Imperfective Future, which contains the Infinitive as one of its components, is represented together with other uses of the Infinitive in Janda & Lyashevskaya 2011. [↑](#footnote-ref-11)
12. Čujkova 2017 presents the results of multiple searches for Future tense forms in the manually disambiguated subcorpus of the Russian National Corpus, showing that the Perfective Future is between four (for the entire subcorpus) and twenty (for the oral portion of the subcorpus) times more frequent than the periphrastic Imperfective Future. [↑](#footnote-ref-12)
13. For example (4): chi-square = 14.88, df = 2, p = 0.0006, Cramer’s V = 0.3. For example (5): chi-square = 3.42, df = 2, p = 0.181. [↑](#footnote-ref-13)
14. The test items in our experiment have more triggers associated with them (17%) than observed for verbs in a corpus (2% according to Reynolds 2016). This is likely an artifact of the fact that several types of verb forms had to be excluded from the experiment (see list in 2.1), and possibly the selection criteria for the texts used in the experiment (again see 2.1). [↑](#footnote-ref-14)
15. For example (8): chi-square = 35.67, df = 2, p = 1.8e-8, Cramer’s V = 0.5. [↑](#footnote-ref-15)