Time and space in parallel streams: in place of an introduction

Laura A. Janda
Your article is protected by copyright and all rights are held exclusively by Springer Science +Business Media Dordrecht. This e-offprint is for personal use only and shall not be self-archived in electronic repositories. If you wish to self-archive your article, please use the accepted manuscript version for posting on your own website. You may further deposit the accepted manuscript version in any repository, provided it is only made publicly available 12 months after official publication or later and provided acknowledgement is given to the original source of publication and a link is inserted to the published article on Springer’s website. The link must be accompanied by the following text: “The final publication is available at link.springer.com”.

Springer
Time and space in parallel streams: in place of an introduction
Время и пространство в параллельных потоках: вместо предисловия

Laura A. Janda

Published online: 27 September 2013
© Springer Science+Business Media Dordrecht 2013

We are pleased to present Space and Time in Russian Temporal Expressions as a special issue of Russian Linguistics, guest edited by Laura A. Janda, Stephen M. Dickey, and Tore Nesset. Here we offer some of the research results of the CLEAR (Cognitive Linguistics: Empirical Approaches to Russian) group1 at the University of Tromsø and our collaborators. This research was sponsored by a grant from the Norwegian Research Council for a project entitled Neat Theories, Messy Realities and by a grant from the Centre for Advanced Study at the Norwegian Academy of Science and Letters for a project entitled Time is Space: Unconscious Models and Conscious Acts.2

The theme of this special issue of Russian Linguistics is the expression of time in Russian, a topic that raises two larger issues in linguistics:

1. What is the relationship between time and space in language?
2. How are choices made between rival linguistic forms?

The first question connects us to the metaphorical structure of linguistic cognition, which draws parallels between the experience of space and the understanding of time. While the overwhelming strategy is to map structures known to exist in the spatial dimension to the temporal one, the relationship is not one-to-one, but also shows divergence between the two domains. The articles in this special issue detail this complex relationship as revealed by adverbs, prepositions, aspectual morphology, and the etymologies of lexical items.

The second question connects us to language structure and processing. Rejecting traditions that have assumed elaborate hierarchical structures in language, a recent proposal (Frank et al. 2012) offers a model based instead on sequential structure. In this model, the constructions of a language appear in parallel, and complex sequences are aggregated by

L.A. Janda (✉)
HSL fakultet, University of Tromsø, 9037 Tromsø, Norway
e-mail: laura.janda@uit.no

2 http://www.cas.uio.no/research/1112timeisspace/index.php.
navigating among parallel streams. This model can account for the choice of rival forms as the selection of elements from one stream while bypassing nearly-synonymous elements in a parallel stream. The contributors to this special issue present data on choices between various lexical items and between grammatical constructions at both the syntactic and the morphological level.

All of the authors represented herein take an empirical approach to their respective topics, relying on various combinations of corpus data and statistical analyses. Some of the corpora and statistical applications described here are highly innovative. Their presentation is designed to promote best practices and encourage other scholars to take advantage of similar methods. This raises a third general theme for this volume:

3. What are the best practices for the empirical investigation of language?

The three sections of this introductory article examine the contributions in this special issue in more detail from the three vantage points outlined above: parallels between time and space, the parallel stream model for rival forms, and the use of new empirical methods to collect and analyze linguistic data. In this introductory article, the contributions in this issue will be referenced by their authors’ names as Nesset et al., Baayen et al., Makarova and Nesset, Kuznetsova et al., and Plungian and Rakhilina.

1 Parallels between time and space and where they break down

All of the articles in this special issue test the limits of the time is space metaphor. While there is plenty of confirmation for the hypothesis that time is understood in terms of space, we find that the metaphor is not sufficient to explain all the phenomena we observe. Time is not merely a temporal version of space. Some structures that are found in the spatial domain do not appear in the temporal one, and the reverse is also true in some cases, suggesting that time can operate independently of space too.

Nesset et al. compare the Russian spatial terms for ‘here’, zdes’ and tut, and find that the latter, but not the former, is often used instead to mark a point as metaphorically ‘here’ in the temporal domain, as in vot ja za stolom, otodvigaju jaščik i tut zamečaju na stole anketu ‘so I am sitting at my desk and I open the drawer and here / now I notice a questionnaire on the desk’. Tut presents a clear case of the time is space metaphor, but there is a restriction that prevents us from applying this metaphor to zdes’. There is likewise a difference between the meanings of ‘now’ conveyed by sejčas and teper’, a difference that depends on a contrast in temporal location, but that does not seem to be motivated by any parallel in spatial expression. Vot (like the French) ‘voilà’ behaves as a verbal pointing gesture incorporating both spatial and temporal deixis without giving primacy to either domain. Nesset et al. conclude that while temporal language does depend on spatial language, time also exercises considerable autonomy from space, leading to space-time asymmetries, which are also identified by Makarova and Nesset.

Makarova and Nesset explore the use of prepositions meaning ‘in(to)’ with temporal nouns in five Slavic languages, as in Russian v ponedel’nik ‘on Monday’ (with the accusative case) and v janvare ‘in January’ (with the locative case). In all such expressions it is reasonable to assume that there is a connection between the (primary) spatial expression of the preposition and its use in temporal adverbials. However, there are also important asymmetries between time and space. In all five languages we observe a complementary distribution of the accusative vs. locative case when ‘in’ is used with temporal nouns, as in the Russian examples above: days of the week require the accusative, but months require the locative. In
the spatial domain the distribution of cases is instead contrastive, distinguishing motion by the use of the accusative from rest by the use of the locative. Case usage in temporal expressions is more limited, but it is also true that temporal constructions, once established, may undergo relatively independent development.

Kuznetsova et al. continue the theme of prepositional constructions in which spatial meanings are extended metaphorically to the domain of time. Pod ‘under’ can signal simultaneity in son pod šum dožđa ‘sleep to the patter of rain’, proximity to a temporal boundary in pod utro ‘towards morning’, and prospectivity in bočka pod kvas ‘barrel for kvass’. Only the last of these meanings is reflected by iz-pod ‘from under’, which can be used to signal retrospectivity in bočka iz-pod kvasa ‘empty kvass barrel’. All of these can be understood as temporal uses, but they differ from the spatial uses of the same prepositions in that very different semantic restrictions are relevant for the noun phrases in the spatial vs. temporal uses. The spatial uses also include specialized mini-constructions that have no correlates in the domain of time. Overall, temporal uses tend to change the relationship between central and peripheral meanings for a preposition, since meanings that are peripheral in the spatial domain often function as central in the temporal domain. Similar to Makarova and Nesset, Kuznetsova et al. found that in temporal expressions prepositions have a reduced range of case use, since pod can combine with both the accusative and the instrumental in spatial expressions, but only with the accusative in temporal ones. The authors attribute the preference of temporal expressions for the accusative case, which indicates motion toward a goal, to the fact that time itself is often perceived as moving.

Russian adjectives denoting speed give Plungian and Rakhilina another perspective on the time is space metaphor. Rather than referring to distance traveled, adjectives like bystryj ‘quick’ are relatively detached from spatial reference and instead compare the temporal properties of a target situation with a standard norm. Physics defines speed as distance divided by time, so a high speed implies covering a large distance in short time, but the same is not true for language. For example, bystryj razgovor ‘quick conversation’ does not indicate that a lot was conversed about in a short time, but rather just that the action was short, and thus this collocation is nearly synonymous with korotkij razgovor ‘short conversation’. Linguistic speed is more about duration than about (metaphorical) distance, and consequently relatively independent of space. There is also a bigger focus on high speed than on low speed in language, since we see more lexical elaborations for the former than for the latter.

The authors of the four articles described above all observe space-time asymmetries, which begs the question as to why such asymmetries should exist. If the conceptual construction of time is a metaphorical mapping of structures from space to time, then we would expect some differences due to the fact that such mappings are typically partial rather than complete (Janda 2010). In other words, we would not expect all of the structures that are found in the domain of space to be mapped over to the domain of time. This would account for greater restrictions in the use of expressions for time as opposed to space, as reported by Makarova and Nesset and Kuznetsova et al. But this does not account for the structures found only in temporal expressions observed in all four articles. This finding deserves more theoretical and empirical consideration in search of alternative motives beyond the time is space metaphor.

2 Activating parallel streams and switching between them

Frank et al. (2012) challenge the prevailing assumption that languages possess a hierarchical structure as modeled in the hierarchical trees posited in generativism. On the basis of a review
of recent neurophysiological, behavioral and computational studies, Frank et al. argue that linguists and psychologists are not justified in assuming hierarchical structure in language processing. They present evidence that sequential structure alone has sufficient explanatory power to account for many, perhaps most, linguistic phenomena. In other words, instead of analyzing utterances as composed of successively more complex units as in (1), it is probably the case that hearers analyze them according to the linear order of components as in (2).

(1) [Sentences [[can [be analyzed]] [as [hierarchically structured]]]]

(2) [Sentences] [can be analyzed] [as hierarchically structured]

Frank et al. (2012) point to a number of problems with the assumption of hierarchical structure in language. For one thing, hierarchical structure has been attributed to an ability that is unique to human beings, but there is neither any viable evolutionary explanation nor any independent evidence for such a unique ability. Furthermore, hierarchical combination is probably too cognitively demanding to be applied recursively as has been asserted. Recent studies in the fields of cognitive neuroscience, psycholinguistics, and computational modeling point instead to sequential structure as the fundamental strategy for the structure of language.

Sequential structure makes sense from the perspective of both how language is used and what we know about how the brain works. Language production and comprehension are necessarily sequential because they take place along the temporal dimension. Sequential structure is well established in terms of evolutionary continuity and general neural mechanisms. In other words, sequential structure does not force us to accept any ad-hoc assumptions that would apply only to human language.

Frank et al. (2012) offer an alternative non-hierarchical model in which components are combined from parallel streams, creating complexity without recourse to hierarchical structure. This model makes room for interactions between components in parallel streams, since various factors can control the switching between them. Sequential processing thus involves the selection of components from parallel streams, and this is relevant because choice between alternative forms is a major theme of this special issue.

Baayen et al. present four different case studies where Russian offers binary choices between forms. At least two of these choices could be described as allomorphy, namely the choice between *nu*-suffixed and unsuffixed forms like soxnuvšij [participle] vs. sox [past tense] ‘dried’ and the choice between the prefix variants *o*- and *ob*- as in osložnit’ ‘make complicated’ [before an obstruent] vs. obnovit’ ‘renew’ [before a sonorant]. However, the data shows that the supposed constraints on allomorphy leak; instead of complementary distribution, there is overlap and competition in some contexts, so one can find both sox and soxnul ‘dried’ and both obskakat’ [with *ob*- before an obstruent] and onemečit’ ‘germanify’ [with *o*- before a sonorant]. There is thus a continuum ranging from allomorphy to rival forms. Even at the allomorphy end of this continuum we find some overlap in certain environments. Rival forms can be both partly overlapping and partly contrastive in distribution, like the prefixes *pere*- and *pre*- and the locative alternation constructions with theme–object in gruzit’ seno na telegu ‘load hay onto the cart’ vs. goal–object in *gruzit* telegu senom ‘load the cart with hay’. A wide variety of factors such as semantics and frequency come into play in regulating the distribution of each of these choices.

Two contributions in this issue discuss rival forms involving case and preposition constructions. Makarova and Nesset examine adverbials expressing when an event takes place across various Slavic languages. In Russian, for example, there are choices between *v* ‘ín(to)’ in construction with either the accusative case as in *v poneděl’nik* ‘on Monday’ or locative
case as in v janvare ‘in January’, plus other competing constructions such as the bare genitive with dates as in pjas naja ‘on the fifth of May’, bare instrumental as in letom ‘in the summer’, and o ‘about’ with the locative as in o Roždestve ‘at Christmas time’. These various options show different distributions across the Slavic languages, and in each language are associated with differences in the number and meaning of the temporal noun and whether it is modified by an adjective. Kuznetsova et al. discuss the competition between constructions such as pod konec vs. k koncu ‘toward the end’ and (pristrojka) pod bassejn vs. dlja bassejna ‘(extension) for a swimming pool’ and find that in both, the version with pod and the accusative case is much more restricted in terms of the nouns that can be used.

The remaining two contributions focus on lexical choices. Nesset et al. find that the five Russian words used to signal ‘here’ and ‘now’, zdes’, tut, sejčas, teper’, vot, inhabit a single network of spatio-temporal meanings, though their centers of gravity and ranges within that network vary. These words show different behaviors with respect both to their overall corpus distribution and to their use in TV broadcasting. Plungian and Rakhilina point out that Russian has many words that express speed, particularly among adjectives that can mean ‘quick’ such as bystryj, skoryj, šustryj, sporyj, etc. They carefully pick apart these near-synonyms, showing how metaphor and metonymy motivate differences in their meanings and use.

What is going on when speakers select one near-synonym out of multiple options for expressing the same (or nearly the same) idea? I will approach this question by examining first form-to-meaning relationships and then possible mechanisms for activating and making selections.

The expressions of a language are pairings of form and meaning, known as ‘constructions’ (Goldberg 2006) or as ‘symbolic assemblies’ (Langacker 2008, pp. 21–24). The latter concept is broad enough to account for all structures of grammar, but here we will use the term constructions to refer to linguistic elements at various levels, including morphemes (like prefixes and suffixes), grammatical constructions (like preposition and case constructions and the locative alternation constructions), and lexical items (like near-synonyms). The form-meaning relationship described here has its roots in De Saussure’s (1949[1916]) concept of language as a system of signs that express ideas.

Ideally one might imagine a one-to-one relationship between form and meaning in a construction, as in Fig. 1.

The form end of the relationship can be complex in the case of allomorphy, where we have a default allomorph and its variants. Similarly, on the meaning end, we find that the majority of linguistic elements are polysemous (Langacker 2008, p. 37). We can represent polysemy as a radial category with the prototypical meaning of the element and its more peripheral meanings. We can diagram a form-meaning relationship with both allomorphy and polysemy as in Fig. 2. Here the various allomorphs are all labeled ‘form’. ‘Form b’ is the default allomorph and ‘meaning b’ is the prototypical meaning. One line connects the entire group of forms to the entire group of meanings in Fig. 2, though the relationships may be more complex, as shown below in Fig. 3.

As shown by Baayen et al. and discussed immediately above, there is no clear division between allomorphic variation and rival forms, and even cases of allomorphy may involve some overlapping contexts. Rather than a strict allocation of forms according to contexts, it is more accurate to talk about statistical tendencies. Both allomorphy and rival forms can be understood as choices among forms that languages offer their users. Perfect synonymy
is likewise quite rare, so choices are further complicated on the meaning end. Rival forms present choices that have overlapping but not necessarily completely identical meanings. What is a prototypical meaning for one near-synonym may be peripheral for another. We can diagram a linguistic choice as in Fig. 3, where solid lines represent stronger connections and dotted lines represent weaker connections (these have been assigned randomly). Both the lines connecting the three forms and the lines connecting the three meanings have been removed.

In Fig. 3, Form b is associated with all three meanings, but more strongly with meanings b and c than meaning a. Form a is most strongly associated with meaning b, though meaning a also occurs, and form c is associated with both meaning b and meaning c, but more so with c. Figure 3 can be adjusted to fit any of the analyses found in this special issue. For example, the Russian words for ‘here’ and ‘now’ (Nesnet et al.) can be diagrammed as in Fig. 4.

Zdes’ connects only to the spatial meaning ‘here’ and vot connects only to the discourse situation ‘here’ and ‘now’. Tut connects to both of those meanings and can also express the temporal ‘now’. Sejčas connects to all the meanings containing ‘now’, but only weakly to uses where there is temporal contrast, which is the only meaning where teper’ is preferred.

Where meaning is held constant, we can substitute context, with the understanding that of course context and meaning are closely connected. In other words, in the absence of any (lexical) meaning difference, different forms (allomorphs) can index different contexts. For example, -u/-ju, -e, and -i are allomorphs of the dative singular ending for Russian nouns. Though the meaning is the same, they do index different contexts since -u/-ju means that the noun is masculine or neuter, -e means that the noun ends in -a in the nominative, and -i means that the noun is feminine ending with a soft sign (mjagkij znak) in the nominative. In a similar, though less clear-cut fashion, the nu-suffix tends to index participles and gerunds, while the Ø-suffix tends to index past tense forms among verbs that show nu- vs. Ø-variation. The distribution of nu- vs. Ø-suffixes in Russian verbs in Baayen et al. could be represented as in Fig. 5.
Both *nu* and *Ø* are found in all three contexts, but whereas *nu* is more strongly associated with the non-finite participle and gerund forms, *Ø* is preferred for finite past tense forms.

Let us assume that Figs. 4 and 5 illustrate our model for the form-meaning relationships, and now tackle the issue of how alternative forms and their meanings can be made available and selected. Here I propose that this can be modeled as selection from parallel streams. This proposal is an extension of the one made by Frank et al. (2012), showing how syntactic relationships that have previously been explained in terms of hierarchical structure can instead be accounted for in a model using sequential, non-hierarchical structure. In their model, items are combined from parallel sequential streams, eliminating any need for hierarchical processing. They give as an example the complex construction *put your knife and fork down*, which includes elements from the more schematic constructions *put X down* and *your X*, plus a chunked phrase *knife and fork* (for more on linguistic chunks, see Dąbrowska 2004, Chapter 9). The whole is processed by ‘changing lanes’ among the parallel streams, as diagrammed in Fig. 6 (note that the vertical ordering of constructions is arbitrary).

This model can be compared to a musical score for a symphony in which different instruments represent different parallel streams. Each instrument is silent for a while and then plays for a certain time, and then stops again. Note also the resemblance to non-linear models of phonology, where each articulator has its own stream.

The parallel sequential streams model proposed by Frank et al. (2012) is primarily aimed at comprehension and at syntax, but there is in principle no reason why it cannot be extended to production data (such as what we find in a corpus) and to other kinds of constructions (such as morphemes, preposition and case constructions, and lexical items). Furthermore, there is no reason why parallel streams should not include alternative rival forms in addition to the forms that are actually selected. In other words, relevant rival forms could be activated as parallel streams without necessarily being selected, and the selection could be regulated by various factors such as frequency, meaning, etc. (Frank p.c. 2013). Indeed we already know from priming studies that exposure to a given linguistic form activates other forms with similar meanings, function, and form, so the idea that the choice of rival forms could involve the activation of parallel streams, only one of which is actually chosen, is plausible.

Figure 7 shows how this extension of the parallel sequential streams model might work for rival forms in the production of the sentence *Avraam nagruzil drova na oslika* ‘Abraham loaded wood onto a donkey’, an example of the use of *gruzit* ‘load’ in the theme-object construction; cf. statistical models of how this verb and construction interact in Baayen et al. (this volume, Sect. 3.1). This sentence illustrates the transitive active verb construction with a subject in the nominative case and an object in the accusative case (Subject-N Verb
Object-A). Russian offers rival forms that can be modeled as parallel streams for both the locative alternation, with the theme–object vs. goal–object constructions, and for the perfectivization of the verb, which can be achieved using three different prefixes. All three perfectivizing prefixes and both types of object constructions can be activated in parallel streams and various factors are responsible for regulating the switching among streams that yields the actual utterance.

A similar model can be used to capture the relationships among rival forms described in all five articles in this issue. Collectively one could say that this issue describes the control mechanisms for switching among parallel streams and selecting one rival form over another. These control mechanisms receive both qualitative and quantitative attention in our articles, as described in the following section.

3 New empirical approaches

All of the studies in this issue rely on authentic language data from corpora, namely the Russian National Corpus (all five articles), the NewsScape Library of International Television News at UCLA (Nesset et al.), and the ParaSol corpus and Czech National Corpus (Makarova and Nesset). It is no exaggeration to state that corpus data have become a mainstay of modern linguistic analysis, regardless of whether or not statistical methods are applied. In addition to providing concrete data for Russian, parallel corpora facilitate cross-linguistic comparisons, as we see in the comparison of temporal adverbials across Slavic languages made by Makarova and Nesset and in the lexical typological comparisons made by Plungian and Rakhilina. Additionally, a multi-modal corpus makes it possible to assess the relationship of linguistic forms to simultaneous visual signals such as gestures and images, as shown by Nesset et al.

The sheer quantity of data is constantly increasing; indeed the Russian National Corpus has doubled in size even during the one year in which this special issue was under production. The NewsScape Library was initiated while our research was underway and already contains billions of words in addition to accompanying images. In the face of so much data, it is becoming more and more common to apply statistical methods in order to probe the data for patterns. Three of the articles in this issue, Baayen et al., Nesset et al., and Makarova and Nesset, apply statistical tests to corpus data. However, this should not suggest any priority of quantitative methods over qualitative ones. The truth is that we need to continue to develop both kinds of analysis in order to move the field of linguistics forward. Without qualitative insights into linguistic questions, quantitative studies are meaningless. One needs considerable qualitative expertise in order to recognize what kinds of questions are linguistically interesting, figure out what kind of data it makes sense to collect, and then interpret the results.
However, even qualitative study should not take place in a vacuum based on armchair mus-
ingings over constructed examples. Instead, qualitative analysis should be informed by corpus data, as illustrated by Kuznetsova et al. and Plungian and Rakhilina.

Whereas qualitative analysis has a long tradition in linguistics, quantitative analysis is less established. We see a need to develop standards, both in terms of what kinds of statistical models to apply to given linguistic questions, and also how to present and share data and code.

The three articles in this issue that apply statistical methods attempt to provide good examples of best practices. Nesset et al. demonstrate differences in the distribution of deictic words in the Russian National Corpus, and each time back up their data with a chi-square test for statistical significance and a test of the effect size. Taken together those two tests tell us whether an observed difference is likely to be indicative of a real difference (rather than being an artifact of a given sample) and whether the difference is big enough to be meaningful. It is reasonable to expect that all studies of corpus data that claim to find different distributions of forms should at minimum present the results of these or similar tests. Makarova and Nesset use some more sophisticated methods, namely a principle components analysis and linear regression to explore the relationship between the grammatical constructions of time adver-
bials and the geographic locations of languages. These two methods yield similar results, confirming an east-west distribution of constructions. Baayen et al. explore the use of classi-
fication tree & forest and naive discriminative learning models as alternatives to traditional logistic regression in modeling the choice of rival forms in Russian. They find remarkable convergence among these three methods in terms of classification accuracy and assessment of the relative importance of variables. Each method has a different set of strengths and weak-
nesses. The tree & forest and naive discriminative learning models are non-parametric, which means that they are probably more appropriate for the kind of distribution of data we often find in language corpora, and they are also more plausible as models of linguistic behavior, since the individual trees ‘vote’ for various choices of rival forms (similar to choices made by individual speakers), and naive discriminative learning adjusts the weights of these asso-
ciations (similar to what we presume happens in neural networks). Baayen et al. conclude that we can recommend tree & forest and naive discriminative learning models as methods to complement logistic models.

The authors of all three studies have posted their data and code to a publicly accessible website at http://ansatte.uin.no/laura.janda/RLdata/RLdata.html and use the open-source R software for their analysis. Readers are encouraged to inspect the data and run the code on their own computers. Our research adheres to the scientific standard of replicability in that anyone can verify our results. We hope that other scholars will complete the cycle of the scientific method by also reproducing similar results for similar kinds of data and research questions.3

4 Concluding remarks

On the level of specific contributions, this special issue offers detailed studies of Russian morphology, syntax, and lexicon used in the expression of time. Beyond that, this issue of Russian Linguistics addresses some core issues in linguistics. It provides a more nuanced

3Collins (1985, p. 19) refers to replicability of results as the “Supreme Court of the scientific system”. Reproducibility is presented as a cornerstone of the scientific method in textbooks like that of O’Leary (2004, pp. 58–64).
perspective on the time is space metaphor, showing that while the expression of time is largely motivated by patterns of spatial expression, time in language does not simply mirror space; time presents different restrictions and can be partially independent of space. The choice of linguistic forms also emerges as a major theme in this issue. Rival forms can be modeled in terms of parallel sequential streams, elaborating on a proposal that minimizes the need for hierarchical structure in language. Data analysis is a growing challenge for linguists, and we suggest some best practices for implementation of statistical methods and public access to files and code. We hope that these studies can serve as models that will inspire others to conduct similar research.

References

Plungian, V., & Rakhilina, E. (this volume). TIME AND SPEED: Where do speed adjectives come from?