A Curious Case of Allomorphy: Russian Verbs Meaning ‘Do It Once’

Laura A. Janda
laura.janda@uit.no
http://hum.uit.no/lajanda
Main Idea

Linguistic definitions tend to be absolute.

Linguistic phenomena tend to be gradient.
In Janda 2007 I presented my “cluster model” of Russian aspect.

Cluster model claims that semelfactive verbs with suffix -nu (like čixnut’ ‘sneeze once’) and with prefix s- (primarily formed from motion verbs, such as sxodit’ ‘walk someplace and come back once’) form a single group of verbs: “Single Act Perfectives”.

Čixnet
“She’ll sneeze once”
But...

• This is a strange distribution between -nu and s-.
My solution

• Allomorphy hypothesis:
  – -\textit{nu} and \textit{s}- serve (approximately) as allomorphs in the formation of semelfactive verbs in Russian
Allomorphy

• Allomorphs are traditionally defined as a group of two (or more) morphemes that have the same function, yet are in complementary distribution (Bloomfield 1935: Chapters 10 & 13; Matthews 1974: Chapter V)

• (usually these are morphemes that are etymologically related but have undergone sound changes in complementary environments)
Examples of allomorphy

• allomorphs of the root *knig-* ‘book’ in the following forms which differ in their final consonants: *kniga* [kn’ig-] (Nsg), *knige* [kn’ig’-] (Lsg), *knig* [kn’ik-] (Gpl), *knižka* [kn’iš- ] (dim Nsg), *knižek* [kn’iž-] (dim Gpl)
  – final segment of morpheme can be g, g’, k, š or ž

• allomorphs of past tense marker: *pisal* [l], *pisala* [l], *pisali* [l’] ‘he, she, they wrote’; *nes* [Ø], *nesla* [l], *nesli* [l’] ‘he, she, they carried’
  – morpheme can be l, l’ or Ø
My desire to escape from that corner and confirm the cluster model led me to a larger theoretical question:

**What is allomorphy?**

...which led me to an even larger question:

**How do we deal with gradient realities when our linguistic definitions are stated in absolute terms?**
Reality is messy

• Our theories often give us clear definitions, but if we look at corpus data, we are often faced with fuzzy gradience. (Newman 2008)

• Does this mean that our definitions must dissolve into relativity?

• No, but we do need to establish standards that account for messy reality.
Overview

• Cluster model and Single Act Perfectives

• Database of Single Act Perfectives with -nu and s-

• Testing the allomorphy hypothesis
  – Are -nu and s- in complementary distribution?
  – Do -nu and s- have the same function?

• Conclusions and future steps
The cluster model distinguishes four types of perfectives:

– Natural Perfectives
  • pisat’ ‘write’ > napisat’ ‘write’

– Specialized Perfectives
  • rabotat’ ‘work’ > pererabotat’ ‘rework, edit’

– Complex Act Perfectives
  • stonat’ ‘moan’ > postonat’ ‘moan for a while’

– Single Act Perfectives
  • dut’ ‘blow’ > dunut’ ‘blow once’

‘He blew once...’
-nu database

• 296 Imperfective verbs that form -nu semelfactives
  – collected by Anastasia Makarova
  – includes both -nu and -anu semelfactives like pleskat’ ‘splash’ which forms plesnut’ and pleskanut’ ‘splash once’
  – includes both reflexive and non-reflexive verbs like kačat’/kačnut’, kačat’sja/kačnut’sja ‘rock/rock once’
s- database

- 105 Imperfective verbs that form s- semelfactives
  - collected by Laura Janda with help from Anastasia Makarova
  - includes eleven motion verbs such as *xodit’/sxodit’* ‘walk/walk somewhere and come back once’
  - includes both reflexive and non-reflexive verbs such as *lovčit’/slovčit’, lovčit’sja/slovčit’sja* ‘be cunning/do something cunning’

*Sxitril?* ‘Did he just do one sneaky thing?’
Comparison of the databases

- 3 times more -nu than s- semelfactives
- -nu semelfactives tend to have higher frequency
- s- semelfactives dominated by 4 motion verbs: sxodit’, s”ezdit’, sbegat’, sletat’ ‘walk, ride, run, fly someplace and come back once’
- Do not analyze frequency because
  - -nu and s- verbs can be both resultative & semelfactive
  - s- prefixed motion verbs have imperfective homonyms like sxodit’ (s uma) ‘walk down (go crazy)’
Are \(-nu\) and \(s\)- allomorphs?

• Are \(-nu\) and \(s\)- in complementary distribution?
  – Almost: verb classes largely determine the distribution of \(-nu\) and \(s\)-

• Do \(-nu\) and \(s\)- have the same function?
  – Almost: there are some verbs that use the two morphemes synonymously and Isačenko (1960) describes semelfactives formed with \(-nu\) and \(s\)- with the same term: \(odnokratnye\) ‘one-time’
Are \textit{-nu} and \textit{s-} in complementary distribution?

- See handout
- A chi-square test shows that the results are statistically significant:
  - the chi-square value is 257.3 with 5 df
  - the probability that this distribution is the result of mere chance is < 2.2e-16 (statistically = 0)
  - Cramer’s V (effect size) = 0.8 (enormous)
Verb classes that prefer -nu
• -aj
• non-productive 1. conjugation
• -*ě

Verb classes that prefer s-
• -ova
• -i
• -*ěj
Verb classes that prefer *nu

-aj

non-prod
1. conj

-*ě

Zevnul
‘He yawned once’

Liznula
‘She licked once’

Svistnula
‘She whistled once’
Verb classes that prefer **s-**

**-ova**

*Smalodušestvoval*

‘He did one cowardly thing’

**-i**

*Sgrubil!*

‘He did one rude thing!’

**-*ěj**

*Srobela?*

‘Was she shy once?’
Complementary distribution: summary

• The distribution is not perfect, but statistically it is pretty close

• For two suffixes there is a perfect distribution: verbs in the non-productive 1. conjugation use only -nu, and verbs with -*ěj use only s-

• For the other suffixes we see strong tendencies, but there is overlap, especially for verbs with the suffixes -ova and -i
Semantic classes in RNC:
Another measure of complementary distribution?

– Only 269 base verbs that form semelfactives with -nu and 37 with s- are tagged (see handout)
– We see clear tendencies, but lack enough s- data for a statistical analysis
– Morphological and semantic classes are not entirely independent factors
Morphological and Semantic Classes

• sound verbs often have the suffixes -aj (kvakat’/kvaknut’ ‘croak/croak once’), -a (lajat’/lajnut’ ‘bark/bark once’), or -*ě (xrapet’/xrapnut’ ‘snore/snore once’)

• verbs that denote behaviors are often suffixed in -i (glupit’/sglupit’ ‘be dumb/be dumb once’) or -*ěj (licemerničat’/slicemerničat’ ‘be hypocritical/be hypocritical once’)

Do **-nu** and **s**- have the same function?

- Both can mean ‘do X once’
- One verb forms synonyms with both **-nu** and **s**-:
  - `xvastat’/xvast(a)nut’/sxvasat’
    - ‘boast/boast once’
- A couple of verbs can use both **-nu** and **s**- simultaneously:
  - `metat’(sja)/smetnut’(sja), ‘leap sideways/leap sideways once’,
  - `trusit’/struxnut’, ‘be a coward/be a coward once’

`Xvast(a)nul ili sxvastal?` ‘Did he boast once?’
-nu and s- are not identical in function

- With -nu we usually have one cycle from a series of repeated events: čixat’/čixnut’ ‘sneeze/sneeze once’, lizat’/liznut’ ‘lick/lick once’
- With s- we often have something that only happened once malodušestovat’/smalodušestovat’ ‘act like a coward/act like a coward once’
Evaluation of the -nu and s-allomorphy hypothesis

• Are -nu and s- in complementary distribution?
• Do -nu and s- have the same function?
• Is the allomorphy hypothesis confirmed?
• Is the cluster model confirmed?

• Almost.
• Almost.
• Pretty much.
• Pretty much.
An apparent case of allomorphy that should be rejected: the “empty” prefixes

• The “purely perfectivizing” prefixes all have the “same” function and are distributed across different verbs, so they should be allomorphs, right?       **BUT**

• 28% of verbs can take more than one “empty” prefix
• a study of *gruzit* ‘load’ shows significant differences in grammatical constructions (co-authors Sokolova and Lyashevskaia)
Relevant constructions

• Accusative case names the load (theme-object)
  – Acc + na/v + Acc (nagruzit’ jaščiki na teležku ‘load the boxes onto the cart’)
  – Acc (zagruzit’ ugol’ budet problematično ‘it will be difficult to load the coal’)

• Accusative case names the container (goal-object)
  – Acc + Inst (on nagruzil sanki proviziej ‘he loaded the sleds with provisions’)
  – Acc (nagruzili telegi i uexali v gorod ‘they loaded the carts and rode into town’)

• Data comes from Russian National Corpus
About the results...

• They are statistically significant
  – Chi-square = 452.827 (p<0.0001, df=6)
    Cramer’s V = 0.507 (large effect)
• Constructions show that the “empty” prefixes behave differently from one another and do **NOT** deserve to be recognized as allomorphs.
Back to the big questions...

• What constitutes allomorphy?
• Complementary distribution is traditionally considered an absolute criterion for allomorphy.
• But is this expectation realistic given that language phenomena often exhibit gradient characteristics?
The traditional definition...

- was proposed long before the advent of electronic corpora and statistical software
- perhaps should be re-evaluated as a prototype rather than as an absolute criterion
- statistical methods make it possible to establish standards for evaluation of gradient phenomena
Future steps for Allomorphy

• Examine a range of case studies with various combinations of criteria and outcomes
  – uncontroversial cases & minimal deviations
  – cases that deviate from only one criterion
  – cases that deviate from both criteria
  – cases that do and do not merit recognition as allomorphy
Future Steps for Statistical Analysis

• Linguistic profiling: probing the structure of language via corpus data
  – Constructional profiling: grammatical constructions that a unit appears in (cf. Janda & Solovyev 2009)
  – Grammatical profiling: grammatical environments a morpheme is found in (cf. Nesset, Janda & Baayen forthcoming)
  – Semantic profiling: relationship between a morpheme and the semantic tags of the words it occurs with
Beyond allomorphy

• Similar considerations can apply to other definitions, such as **allophony**, **markedness**, and **neutralization**

• Linguistics has lagged behind other fields in **establishing statistical standards**

• Philosophical debate on the balance between **rationality** and **reasonableness** (Toulmin 2001)