Synonymy

Laura A. Janda
Overview

1. Synonymy from the perspective of Cognitive Linguistics
2. Constructional profiling: ‘Happiness’ and ‘sadness’ in Russian
3. Constructional profiling: ‘Load’ verbs in Russian
5. Radial category profiling: Two ways to get out in Russian
1. Synonymy from the perspective of Cognitive Linguistics

If all linguistic phenomena are motivated by meaning...

...true synonymy is not expected to be widespread in language.

In other words, difference in form is likely to be motivated by difference in meaning.

- Near synonyms ("rival forms") are usually distinguished by their behavior.
- Behavior refers to co-occurrence patterns, association with various factors.
- Behavior can be investigated by means of "profiling".
2. Constructional profiling: ‘Happiness’ and ‘sadness’ in Russian
Constructional Profiles of Nouns

- A constructional profile is “the distribution of relative frequencies of constructions associated with a given word”
- There are about 70 constructions of the form “[(preposition) [noun]_{case}]” in Russian
- The null hypothesis is that all nouns should have equal frequency in all constructions
- But only about 6 (or fewer) constructions are needed to characterize a given noun
- Example of constructional profile of восторг
Where the data come from

- Russian National Corpus (http://www.ruscorpora.ru) >120M words
- Biblioteka Maksima Moškova (http://lib.ru/) >600M words
- 500 sentences extracted and coded for use of preposition & case
vostorg 'ecstasy'

percent frequency

`v+Acc` `v+Loc` `Inst` `s+Inst` `ot+Gen`

‘into’ ‘in’ agent ‘with’ ‘from’
Synonymy

- Hypothesis: Each word has a unique constructional profile
- Corollary: Words with similar meanings should have similar constructional profiles
- Hierarchical cluster analysis shows which constructional profiles are closest (closer synonyms) and which are further apart, using squared Euclidean distances based on constructional profile data
‘Sadness’ in Russian

- грусть, меланхолия, печаль, тоска, уныние, хандра

- The constructions they appear in:
  - в + Acc ‘into’
  - в + Loc ‘in’
  - Inst: Agent
  - с + Inst ‘with’
  - от + Gen ‘from’
  - (Direct Object)
  - (Other Constructions)

Graphs will show only these five, as percentages

Hierarchical Cluster Analysis is computed from all data
The ‘sadness’ nouns

- They can’t all be the same:
  Уходишь, и я гляжу вслед тебе с грустью, но без тоски.
  ‘You leave and I watch you go with sadness\textsubscript{grust}, but without sadness\textsubscript{toska}.’

- Dictionaries differ on how synonyms are grouped:
  - Most usual grouping: грусть, печаль, тоска vs. меланхолия, уныние, хандра
  - Disagreements over уныние:
    - Apresjan et al. 1997: уныние goes with печаль
    - Aleksandrovna 1989: уныние goes with меланхолия and хандра
    - Evgen’evna 2001: уныние goes with грусть and хандра
    - Švedova 2003: уныние goes with грусть, меланхолия and хандра
‘Sadness’
Hierarchical Cluster

печаль  тоска  хандра  меланхолия  грусть  уныние
‘Happiness’ in Russian

- ликование, наслаждение, радость, счастье, удовольствие, восторг
- Antonyms are words that are virtually identical, but differ in one value
- ‘Happiness’ nouns focus on the same constructions in their constructional profiles as ‘sadness’ nouns
- Dictionaries differ widely in grouping of ‘happiness’ synonyms
  - Aleksandrova 1998: наслаждение & удовольствие vs. радость & ликование vs. восторг
  - Švedova 2003 and Abramov 1994: наслаждение & удовольствие vs. ликование, радость & восторг
‘Happiness’
Hierarchical Cluster
About the results...

- The results are statistically significant
- For ‘sadness’ nouns: chi square = 730.35, and Cramer’s V = 0.305 which qualifies as a moderate effect (p<0.0001, df=30)
- For ‘happiness’ nouns: chi square = 774.6, Cramer’s V = 0.268 which qualifies as a moderate effect (p<0.0001, df=30)
‘Sadness’ & ‘Happiness’
Hierarchical Cluster

Group 1
- pečal'
- likovanie
- toska
- radost'

Group 2
- naslaždenie
- udovol', stvie'
- grust'

Group 3
- melanxolija
- xandra
- vostorg
- unynie

20
group 1 combined cluster

percent frequency

constructions

v+Acc  v+Loc  Inst  s+Inst  ot+Gen

nouns
- pechal'
- likovanie
- toska
- radost'
group 3 combined cluster

percent frequency

constructions

v+Acc  v+Loc  Inst  s+Inst  ot+Gen

nouns
- melanxoliija
- xandra
- vostorg
- unynie
3. Constructional profiling: ‘Load’ verbs in Russian
Constructional Profiles of Verbs

- The distribution of relative frequencies of constructions associated with verbs
- Constructional profiles show that near-synonyms can behave very differently
- Constructional profiles can show that the “empty” prefixes are not really empty
“Empty” prefixes?

When we have aspectual pairs such as *pisat’/napisat’* ‘write’, *morozit’/zamorozit’* ‘freeze’, *obedat’/poobedat’* ‘eat lunch’, it is assumed that the prefixes *na-*, *za-*, *po-* are “empty” (have no meaning)

Some verbs have several “empty” prefixes: *gruzit’* ‘load’ has the perfectives *nagruzit’, zagruzit’, pogruzit’* – are they synonyms?

Constructional profiles show that the verbs have different meanings and the prefixes are not empty
The Locative Alternation in Russian

• **Theme**-object construction
  
  *gruzit’ seno na telegu*
  
  ‘load hay onto the wagon’

• **Goal**-object construction
  
  *gruzit’ telegu senom*
  
  ‘load the wagon with hay’

• Variables: VERB (prefixes), (passive) PARTICIPLE, REDUCED

  VERB: unprefixed *gruzit’* or prefixed: *nagruzit’, zagruzit’, pogruzit*

  PARTICIPLE:

  • **Theme**-object: *seno gruženo na telegu* ‘hay is loaded onto the wagon’
  
  • **Goal**-object: *telega gružena senom* ‘the wagon is loaded with hay’

  REDUCED:

  • **Theme**-object: *gruzit’ seno* ‘load the hay’
  
  • **Goal**-object: *gruzit’ telegu* ‘load the wagon’
All non-passive forms

Past passive participles
The Locative Alternation in Russian

**RIVAL FORMS:** the two constructions, theme-object vs. goal-object

**DEPENDENT VARIABLE:**

CONSTRUCTION: theme-object vs. goal-object

**INDEPENDENT VARIABLES:**

**VERB:** zero (for the unprefixed verb `gruzit`) vs. `na-` vs. `za-` vs. `po-`

**PARTICIPLE:** yes vs. no

**REDUCED:** yes vs. no

**DATA:** 1920 sentences from the Russian National Corpus, only 1 example per author
About the results...

• Logistic regression shows a highly significant relationship between the construction and the following factors:
  – prefix
  – passive vs. active
  – reduced vs. full
• Correlational strength is 0.796
• Classificatory power: 88.5 % constructions are predicted correctly
“Ways of trying in Russian: clustering behavioral profiles”

Divjak, Dagmar and Stefan Th. Gries 2006
What are the problems?

- Let’s look at the entries in synonym dictionaries and think about what the problems are.
- Think also about what we know from cognitive linguistics about meaning and category structure.
‘TRY’

• Апресян et al. 1999
  – 1: пробовать, пытаться, стараться, силиться
    • 2: добиваться, домогаться, хотеть, намереваться, стремиться, рваться, порываться, биться, осилить, тщиться, пыжиться, норовить, испытывать
  – Чернова 1996
    • пробовать, пытаться, стараться, норовить, силиться, тщиться, искать, домогаться, не почесаться
  – Евгеньева 2001
    • Пытаться: пробовать
    • Стараться: стремиться, пытаться, норовить, силиться, тщиться, пыжиться
What are the problems?

- Where does a category/semantic space begin/end? Is that a relevant question?
- How can we explore the structure of the category = which meanings are closer to each other, which are further apart?
- Much of the work has been merely intuitive, pairwise judgements
What are the assumptions/methods?

• Distributional similarity = semantic similarity
  – Do we agree?
  – What is distribution?
• Behavioral profile via ID tags
  – How does behavioral profile differ from constructional profile?
• Cluster analysis > radial network
What are the goals?

• Find the structure of the category of tentative verbs
• Find the prototype for each verb
The nine verbs in this study

• пробовать, пытаться, стараться, силиться, норовить, порываться, тщиться, пышиться, тужиться

• Why these nine verbs?
  – All of them share a set of distributional properties...
The distributional properties they share

• All can combine with an infinitive (298 verbs in Russian do this, cf. Divjak’s dissertation)
• None of these verbs can:
  – Кто планировал это сделать?/ Кто планировал это?
    • *Он старался это (Он пробовал писать ≠ Что он пробовал!)
  – Полковник приказал майору занять новые позиции/
    Полковник приказал, чтобы майор занял новые позиции
    • *Полковник старался, чтобы майор занял новые позиции
  – Они постановили вчера завершить строительство завтра
    • *Они старались вчера завершить строительство завтра
What is the behavioral profile?

- ID tagging done manually for all 1585 sentences
- All sentences contained Vfin Vinf construction
- Each sentence tagged for 87 variables (Divjak 2004: 234-237)
  - Aspect, mood, tense of Vfin
  - Aspect of Vinf
  - Structure of clause, subject
  - Adverbial specifications, particles, negation, connectors
  - Apresjan’s primitives (enlarged) -- subjective

How does this compare with a constructional profile?
How behavioral profiles were analyzed

• Each verb produced a spreadsheet with percentages for each of the 87 values
• Hierarchical agglomerative clustering (HAC)
  – 9 verbs = “items”
  – 87 ID tags = “variables”
HAC yields three groups, corresponding to a radial network:

- Three ICMs: YOU WON’T SUCCEED; YOU COULD SUCCEED, YOU CAN’T SUCCEED
  - YOU COULD SUCCEED:
    - Пробовать; пытаться, стараться
  - YOU CAN’T SUCCEED:
    - Тужиться; тщиться, пыжиться
  - YOU WON’T SUCCEED:
    - Силиться; порываться, норовить
More about the three ICMs

• YOU COULD SUCCEED: “a human is exhorted to undertake an attempt to move himself or others (rather than to undertake mental activities); often, these activities are negated”

• YOU CAN’T SUCCEED: “an inanimate subject (concrete or abstract) attempts very intensely but in vain to perform what typically are metaphorical extensions of physical actions”

• YOU WON’T SUCCEED: “an inanimate subject undertook repeated non-intense attempts to exercise physical motion; the actions are often uncontrollable and fail because of internal or external reasons”
The behavioral profile of YOU COULD SUCCEED:
Пробовать; пытаться, стараться

• YOU COULD SUCCEED has high values for:
  – Main clause, infinitive, perfective,
  – Infinitive that follows often negated, expresses physical activities or figurative extensions of motion events or putting an “other” into motion
  – Subject got permission to try (пусть), the attempt was brought to an untimely halt (было), subject was exhorted to attempt, attempt intensity was reduced

• Пробовать is the central verb
The behavioral profile of YOU CAN’T SUCCEED:
Тужиться; тщиться, пыжиться

• YOU CAN’T SUCCEED has high values for:
  – Imperfective (exclusively), present tense
  – Inanimate subjects (concrete & abstract) or groups & institutions
  – Following infinitive refers to physical actions that may affect an “other”, or metaphorical extensions of physical actions
  – Focus on vainness and intensity of a durative effort
YOU WON’T SUCCEED: Силиться; порываться, норовить

- YOU WON’T SUCCEED has high values for:
  - Participles
  - Following infinitive expresses physical motion, often uncontrollable, often repeated
  - Inanimate subject
  - Both internal and external reasons are given in case of failure
Advantages of this approach

• Objective, precise, potentially replicable and reliable
• Confirms radial category structure
5. Radial category profiling: Two ways to get out in Russian

Radial Category Profiling and Russian prefixes: The Case of vy- and iz-

Tore Nesset
Laura A. Janda
Anna Baydimirova
Two issues

• Conventional wisdom 1 – čistovidovye pristavki:
  – Prefixes that form aspectual partners from simplex verbs are semantically empty (aside from changing the aspect to perfective)

• Conventional wisdom 2 – Slavonicisms:
  – Elements of Church Slavic origin have more abstract meanings than elements of East Slavic origin.
  – Is it possible to test these hypotheses empirically?

We propose a method: Radial Category Profiling.
Emptiness vs. Overlap

- *idti* ‘walk’ – *VY-jti* ‘walk out (pf)’
  - Prefix changes verb meaning
  - *VY*-jti = Specialized Perfective
  - *VY*- ≈ ‘out’
- *ryt’* ‘dig’ – *VY-ryt’* ‘dig (pf)’
  - Prefix does NOT change verb meaning
  - *VY-ryt’* = Natural Perfective
  - *VY* = Ø or *VY*- ≈ ‘out’?

Hypothesis:
Prefix bleached to Ø

Hypothesis:
‘Out’ is part of the meaning of *ryt’*

In order to test the Overlap Hypothesis, we need a better understanding of the meaning of *VY-*.
Prediction from Overlap Hyp: Verb stems that form natural perfectives with VY- have meanings attested in this network.

Prediction from Emptiness Hyp: Verb stems that form natural perfectives with VY- have random meanings.
Natural perfectives: Solid lines

Since natural perfectives form a subset of the radial category for VY-in specialized perfectives, we have an argument in favor of the Overlap Hypothesis.
What about \textit{IZ-}? Specialized perfectives

IZ- attested in yellow subcats

1. Out of container
2. Out of metaphoric container
3. Empty container
4. Empty metaphoric container
5. Exhaustive result
6. Exhaust surface
7. Negative exhaustive
8. Make image on surface
9. Make out of
10. Decline/deviate
11. Acquire
12. Overcome
Results favor Overlap Hyp: Verb stems that form natural perfectives with IZ- have meanings attested in the same network as specialized perfectives.
Summary so far

• Four types of verbs in one network:
  – Specialized perfectives in VY-
  – Natural perfectives in VY-
  – Specialized perfectives in IZ-
  – Natural perfectives in IZ-

• Evidence in favor of Overlap Hypothesis
  – Prefixes maintain meaning in all types

• Is the meaning of the prefix identical in all types?
  – No, different subparts of the network are attested for different verb types

Is it possible to sharpen this insight further?
Subcategories have different numbers of members (type frequencies)

Radial Category Profile: The relative frequency distribution of the subcategories of a radial category

Profiles of different categories can be compared using simple statistical methods
Radial Category Profiles for VY- and IZ-

- In order to facilitate statistical analysis we compare three groups of semantically related subcategories
- Pearson’s chi-squared test shows that differences are statistically significant: $p = 6.674e^{-12}$
- Effect size is large: Omega value is $0.483$
Abstractness: VY- vs. IZ-

- Subcategories 1-4:
  - Physical motion out of container landmark
  - Ex: vyjti iz komnaty ‘walk out of the room’
  - Tangible, concrete meaning
- Subcategories 5-7:
  - Emphasize exhaustiveness of change of state
  - Ex: izorvat’ bumažku ‘tear the whole piece of paper into pieces’
  - More abstract meaning
- Prediction:
  - Since IZ- is of Church Slavic origin
  - we expect a more abstract meaning
  - i.e. relatively more members of subcategories 5-7

This prediction is borne out by the facts.
The abstractness scale

• Radial Category Profiling enables us to place the four verb types on an abstractness scale:

```
Concrete  VY_{sp}  IZ_{sp}  VY_{np}  IZ_{np}  Abstract
```

• Generalizations:
  1. $IZ$- is more abstract than $VY$-
  2. For a given prefix, natural perfectives are more abstract than specialized perfectives
  3. Choice of prefix and choice of type of perfective have similar effects
  4. We cannot decide which factor has the stronger impact
Linguistic Profiling in cognitive linguistics

• Behavioral profiles (Divjak & Gries 2006)
  – Reflect corpus frequency of syntactic, morphological and semantic parameters for a word

• Constructional profiles (Janda & Solovyev 2008)
  – Reflect corpus frequency of constructions associated with a given word

• Grammatical profiles (Janda & Lyashevska forthcoming)
  • Reflect corpus frequency of the inflected forms of a word

RCP is another type of linguistic profiling:
Related approaches involving statistical analysis of relative frequencies
Wrapping up

• **Radial Category Profiling**
  – Statistical analysis of relative frequencies of subcategories in radial networks

• **Overlap Hypothesis**: empirical evidence that
  – Even in natural perfectives, prefixes have a meaning
  – In natural perfectives prefixes have more abstract meanings

• **Slavonicisms**: empirical evidence that
  – Elements of Church Slavic origin (e.g. *IZ*) have more abstract meanings than elements of East Slavic origin (e.g. *VV*)
References