



# Favored shifts in derivational morphology accompany expansive contact situations

Johanna Nichols

UC Berkeley

johanna@berkeley.edu

*Grammatical hybridization and social conditions*

Leipzig, October 16-18, 2014

Edited but still a **draft**. Please don't quote or circulate without permission.

## Initial questions

Observation: Diffusion-resistant variables diffuse readily in a few places: eastern Eurasian steppe, north central Caucasus.

Previous work has identified favored selection targets, including causativization (base intransitivity) in causative alternation  
And described the sociolinguistics of frontier situations in language spreads

- Is it really frontier sociolinguistics that leads to expansion of attractors? Or just contact in general?
- How real are these effects?
- Is the favored state of causativization around the eastern steppe really a result of selection? Or just a Pacific Rim macroareal feature?

## Goals of this study

Survey a number of areas with histories of contact

Survey several variables, mostly from the domain of derivational morphology

Hope to produce a respectable pilot study

## Design of survey

Sociolinguistics of expansive contact situations (*frontier conditions*):

Economic expansion; opportunities at edge

Variable, back-and-forth, short-term sociolinguistic dominance

No standard language, minimal or no political authority

No ideology of language discreteness, language identity, purism

Frequently, catalyst language expands at frontier of spread

Known examples: eastern steppe, northeast central Caucasus

Likely further examples: western steppe periphery; catalyst

north of Indo-Aryan spread; maybe catalyst(s) around

northern Kazakhstan, 6000 BP; maybe the northern Pacific

Rim macropopulation

## Design of survey

Areas surveyed: (\* = not complete)

West Africa \*

Balkan

Caucasus; and subparts Avar sphere, eastern Transcaucasus

Western Eurasian steppe periphery

BMAC, early Bronze Age \*

(Bactria-Margiana Archaeological Complex)

Eastern Eurasian steppe (Kazakhstan, Mongolia, nearby)

North Pacific Rim

Northern Australia \*

Western North America; and coastal subpart

Mesoamerica \*

Amazonia \*

Total languages surveyed: 104

## Design of survey

Areas surveyed: (\* = not complete)

West Africa \*

Balkan

*Caucasus; Avar sphere; eastern Transcaucasus*

Western Eurasian steppe periphery

BMAC, early Bronze Age \*

*Eastern Eurasian steppe* (Kazakhstan, Mongolia, nearby)

*North Pacific Rim*

Northern Australia

*Western North America; coastal portion*

Mesoamerica \*

Amazonia \*

Italics = known or inferred frontier-like situation

Fine print: Sociolinguistic synopsis by area. Italics = frontier conditions.

West Africa: Large populations, dominant languages. Areality uncertain.

Balkan area: Known areality. Large populations, ideology of discreteness.

Caucasus: No overall areality. Diversity of residual zone.

*Avar sphere*: Northeast central Caucasus, where Avar was lingua franca. Back-and-forth spreading between Avar and Andi foothills; repeat uphill spreads.

*Eastern Transcaucasus*: eastern Georgia, Armenia, northern Azerbaijan (historical Alwan = Caucasian Albania). Spread zone.

Western Eurasian steppe periphery: Repeat spreading; sociolinguistics unknown.

BMAC, early Bronze Age. Oases = language sumps, countryside = spread zone.

Spreading from countryside, with dominance (Indo-Iranian, then Turkic). Hard to survey; pre-I-I language known only from substratal effects on early I-I.

*Eastern Eurasian steppe* (Kazakhstan, Mongolia, nearby): Spread zone with clear frontier conditions at edges.

*North Pacific Rim* (Japan to central California). Diffuse but clear areality; linguistic symptoms of frontier, though actual sociolinguistics unknown.

Northern Australia: Diversity with ideology of language discreteness.

*Western North America*: Spread zones (Great Basin, Columbia Plateau).

*Coastal western N. America*: Part of North Pacific Rim.

Mesoamerica: Areality due to long-standing contact. History of language dominance.

Early Uto-Aztecan as likely catalyst; expect frontier conditions there.

Amazonia: Diversity with ideology of language discreteness.

## Design of survey

Typological variables surveyed, 1:

Causativization as main realization of causative alternation  
(on 18-pair verb list of Nichols, Peterson, Barnes 2004)



Realization of the causative alternation: Russian vs. Ingush  
 (Valence-deriving morphology in **blue**.)  
 (Ingush: Nakh-Daghestanian, north central Caucasus)

| <i>Gloss</i> | <i>Valence</i> | <i>Russian</i>       | <i>Ingush</i>           |
|--------------|----------------|----------------------|-------------------------|
| 'spoil'      | intrans.       | portit'- <b>sja</b>  | talx                    |
|              | trans.         | portit'              | tolxa- <b>d.u</b>       |
| 'hide'       | intrans.       | prjatat'- <b>sja</b> | dwa-lachq'              |
|              | intrans.       | prjatat'             | dwa-lochq'a- <b>d.u</b> |
| 'fear'       | intrans.       | bojat'- <b>sja</b>   | qier                    |
| 'scare'      | trans.         | pugat'               | qiera- <b>d.u</b>       |

Same semantic relationship (transitive has causative semantics);  
 opposite morphological treatment.

18-pair verb list for surveying realization of causative alternation

|                |            |                   |                  |
|----------------|------------|-------------------|------------------|
| laugh          | make laugh | boil              | boil             |
| die            | kill       | burn (catch fire) | burn (set afire) |
| sit            | seat       | break             | break            |
| eat            | feed       | open              | open             |
| learn          | teach      | be(come) dry      | dry, make dry    |
| see            | show       | straighten        | straighten       |
| be(come) angry | anger      | hang              | hang up          |
| fear           | scare      | turn over         | turn over        |
| hide           | hide       | drop              | fall             |

## Types of derivational pairing in the causative alternation: Examples

|                                |            |   |   |         |
|--------------------------------|------------|---|---|---------|
| Causativize (transitivize)     | Nanai      | inekte-                                   | inekte- <b>we:n-</b>                      | 'laugh' |
| Decausativize (detransitivize) | Russ.      | učit'- <b>sja</b>                         | učit'                                     | 'learn' |
| Double                         | Sib. Yupik | aghagh- <b>nga-</b>                       | aghagh- <b>te-</b>                        | 'hang'  |
| Ambitransitive                 | Ger.       | brechen                                   | brechen                                   | 'break' |
| Light verb construction        | Ingush     | ieghaz- <b>d.uoda</b><br>angry- <b>go</b> | ieghaz- <b>d.ug</b><br>angry- <b>lead</b> | 'angry' |
| Adjective ~ verb               | Engl.      | angry                                     | anger                                     |         |
| Ablaut                         | Russ.      | <b>s</b> idet'                            | <b>s</b> adit'                            | 'sit'   |
|                                | Engl.      | <b>s</b> it                               | <b>s</b> eat                              |         |
|                                | OCS        | <b>vy</b> knŏti                           | <b>u</b> čiti                             | 'learn' |
| Suppletion                     | Engl.      | fear                                      | scare                                     |         |
|                                |            | die                                       | kill                                      |         |

## Design of survey

Typological variables surveyed, 2:

High consistency in realization of causative alternation

Measured as high % in plurality realization type

## Design of survey

Typological variables surveyed, 3:

High frequency of light verb constructions

## Design of survey

Light verb constructions: e.g. 'help' in Nakh-Daghestanian

|                 |                          |                     |
|-----------------|--------------------------|---------------------|
| Ingush, Chechen | gho d.u                  | (help:NOUN make/do) |
| Avar            | kumek habize             |                     |
| Karata          | kumak gaa-lha            |                     |
| Hinuq           | kumak b.uwa              |                     |
| Tsakhur         | kumag ha7as, kumak hiles |                     |
| Lezgi           | kümak awun               |                     |

etc. – **every** Daghestanian language uses Kumyk *kömek* or Azeri *kömek* 'help' (noun) in a light verb construction with 'make/do' or 'give'. Nakh (Ingush, Chechen): different first element, same kind of LVC. But not West Caucasian or Kartvelian.

## Design of survey

Typological variables surveyed, 4:

Noun-based derivational morphology

A noun-based language: Russian. (Nichols, Peterson, Barnes 2004 wordlist, used as conveniently available sample.) Bold: denominal. Italics: Maybe denominal.

|                 | <i>Plain</i>             | <i>Semantic causative</i>             |                                 |
|-----------------|--------------------------|---------------------------------------|---------------------------------|
| <i>laugh</i>    | <b>smejat'-sja</b>       | <b>smešit'</b>                        | < smex 'laughter'               |
| <i>die</i>      | umirat'                  | ubivat'                               |                                 |
| <i>sit down</i> | <b>sadit'sja / sest'</b> | (u-, po-) <b>-sadit' / -usaživat'</b> | (Nichols 2010)                  |
| <i>eat</i>      | est'                     | <b>kormit'</b>                        |                                 |
| <i>learn</i>    | <b>učit'sja</b>          | <b>učit'</b>                          |                                 |
| <i>see</i>      | <i>videt'</i>            | <i>pokazyvat'</i>                     |                                 |
| <i>angry</i>    | <b>serdit'sja</b>        | <b>serdit'</b>                        |                                 |
| <i>afraid</i>   | bojat'sja                | <b>pugat'</b>                         |                                 |
| <i>hide</i>     | <i>prjatat'sja</i>       | <i>prjatat'</i>                       |                                 |
| <i>boil</i>     | kipet'                   | <b>kipjatit'</b>                      | (also <b>varit'sja varit'</b> ) |
| <i>burn</i>     | <i>goret'</i>            | žeč'                                  |                                 |
| <i>break</i>    | <i>lomat'sja</i>         | <i>lomat'</i>                         | (etc.)                          |

Hallmark of noun-based derivation: Many denominal verbs.



A verb-based language: Old Church Slavic (and Late Proto-Slavic)  
(selected examples with Russian cognates)

|              |     | <i>Plain</i> | <i>Semantic causative</i> |
|--------------|-----|--------------|---------------------------|
| <i>die</i>   | OCS | u-mьr-ě-     | u-mor-i-                  |
|              | R   | umeret'      | ubit'                     |
| <i>learn</i> | OCS | vyk-noN-     | uč-i-                     |
|              | R   | učit-'sja    | učit'                     |
| <i>boil</i>  | OCS | vьr-ě-       | var-i-                    |
|              | R   | varit-'sja   | varit'                    |

Hallmark of verb-based type: Many deverbal nouns.  
(Deverbal nouns here are bases for transitives.)

A verb-based language: Thompson (Salish). Examples of semantically basic nouns that are formally deverbal. 7 = glottal stop, 9 = pharyngeal.

|        |                 |  |
|--------|-----------------|--|
| tongue | təʔ-e7          | 'stick out (round flexible object)'      |
| wing   | s-c'k=áXn       | s- nominalizer                           |
| hair   | s-/q'ép=qn      | s- nominalizer                           |
| mother | s-/kíx-ze7      | s- nominalizer                           |
| father | s-/qac-ze7      | s- nominalizer                           |
| moon   | má9=xé-tn       | light:up=foot-INSTR                      |
| sun    | s-k'wák'wes     | s- nominalizer                           |
| dog    | s-/qáXa7        | s- nominalizer                           |
| eagle  | 7es-/kwl-o7=qin | STAT-yellow=head ('head' is lex. suffix) |
| tree   | s-/ghép         | s- nominalizer 'erect (a pole)'          |

(JN and W. A. Foley, ongoing work. Source: Thompson & Thompson 1996.)

## Design of survey

Typological variables surveyed, 5:

High part-of-speech flexibility of derivational base

Format of next slide and several subsequent similar tables:

Entries:

Mean % of verbs (from the 18-pair list) that use causativization  
Standard deviation of that mean

Two columns of verbs, corresponding to the two sets of pairs on slide 10  
above: first set has (prototypically) animate S/O, second set  
inanimate.

Highlighting:

Pink = high (1 s.d. or more above mean)

Blue = low (1 s.d. or more below the mean)

Interpretation:

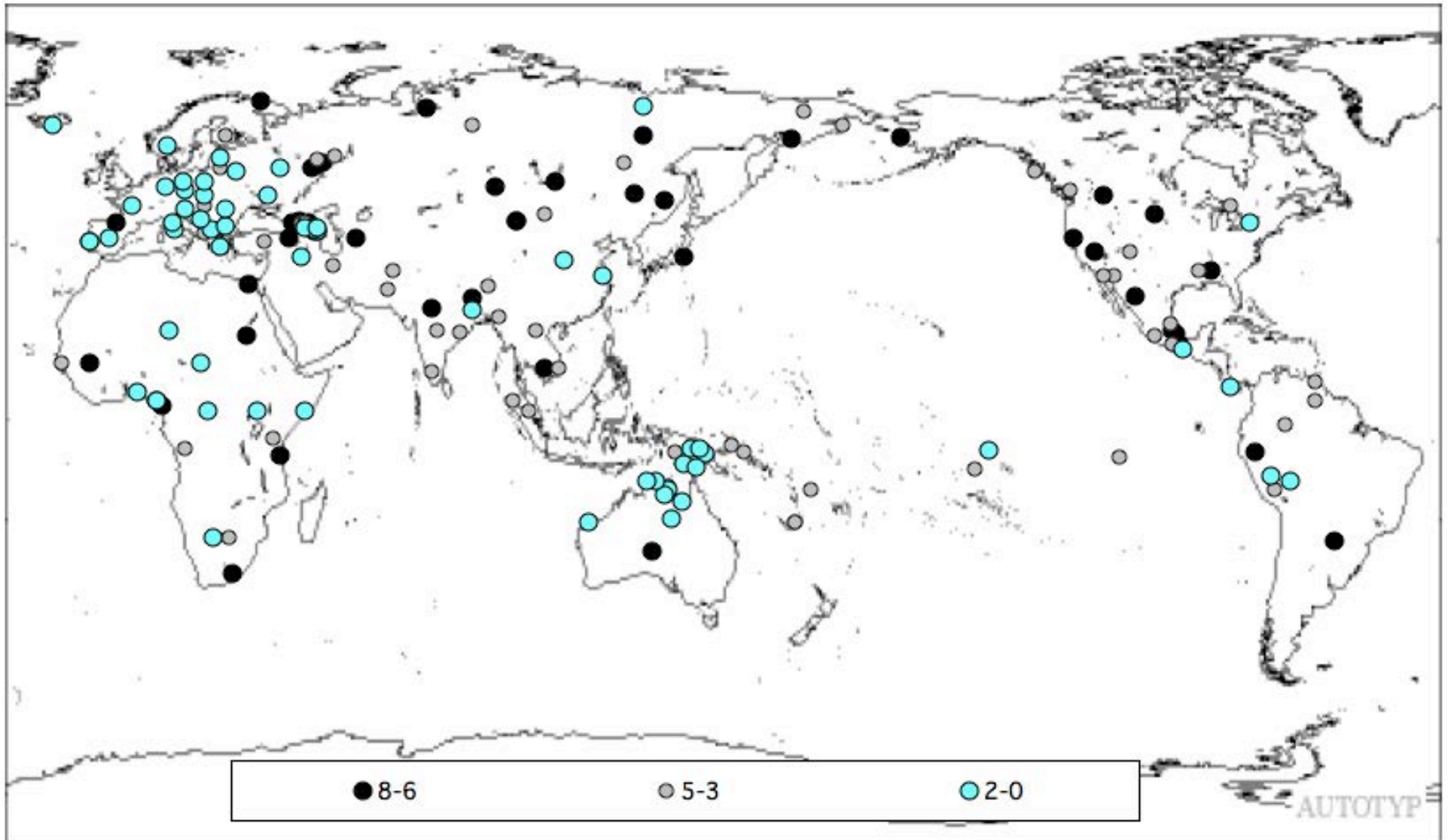
High entries confirm the hypothesis

Low standard deviations indicate areality.

## Findings: Causativization in causative alternation

|                          | Animate verbs |      | Inanimate verbs |      |
|--------------------------|---------------|------|-----------------|------|
|                          | Mean          |      | Mean            |      |
|                          | % caus.       | s.d. | % caus.         | s.d. |
| W. Africa                | 0.44          | 0.27 | 0.36            | 0.33 |
| Balkan                   | 0.06          | 0.06 | 0.02            | 0.05 |
| Caucasus                 | 0.46          | 0.25 | 0.37            | 0.26 |
| E. Transcaucasus         | 0.51          | 0.15 | 0.36            | 0.16 |
| W. steppe periphery      | 0.24          | 0.24 | 0.24            | 0.18 |
| BMAC (Indica)            | 0.51          | 0.23 | 0.49            | 0.21 |
| E. steppe periphery      | 0.68          | 0.16 | 0.50            | 0.18 |
| N. Australia             | 0.17          | 0.19 | 0.33            | 0.36 |
| W. N America             | 0.64          | 0.15 | 0.59            | 0.27 |
| Mesoamerica              | 0.53          | 0.18 | 0.29            | 0.16 |
| Amazonia                 | 0.28          | 0.24 | 0.55            | 0.25 |
|                          |               |      |                 |      |
| Mean ± 1 s.d.            | 0.65          | 0.13 | 0.54            | 0.13 |
|                          |               |      |                 |      |
| Avar sphere              | 0.66          | 0.16 | 0.56            | 0.19 |
| Other N-D                | 0.29          | 0.18 | 0.28            | 0.23 |
|                          |               |      |                 |      |
| E. steppe near periphery | 0.80          | 0.07 | 0.53            | 0.19 |
| E. steppe far periphery  | 0.57          | 0.15 | 0.47            | 0.19 |
|                          |               |      |                 |      |
| N. Am. west coast        | 0.68          | 0.14 | 0.68            | 0.22 |
| Intermontane             | 0.61          | 0.16 | 0.52            | 0.30 |

Number of causativizing pairs (animate verbs)



Comments on preceding map:

Old map, doesn't include the last several languages surveyed  
(chiefly Caucasus, Africa, North America)

Take-home message:

Large compact clusters of languages with minimal causativization in Europe,  
Australia-New Guinea, maybe Africa.

Local clusters of high causativization in Caucasus (west, central), eastern  
steppe and its periphery, north Pacific Rim.

These are spread zones or (Caucasus) their edges, making it look like  
causativization is universally favored.

## Findings: Causativization in causative alternation

### Hypotheses:

1. All contact situations favor high % of causativizing pairs
2. Expansive contact situations favor high % causativizing pairs

### Results:

1. **No.** Counterexamples:  
Balkan (6%), northern Australia (17%), in part Amazonia (28%)
2. **Yes.**  
Eastern steppe (68%), especially near periphery (80%)  
Avar sphere (66%)  
North American west coast (68%)



## Findings: High consistency in causative alternation

|                          | Animate verbs  |      | Inanimate verbs |      |
|--------------------------|----------------|------|-----------------|------|
|                          | % in plurality | s.d. | % in plurality  | s.d. |
| W. Africa                | 0.55           | 0.16 | 0.54            | 0.18 |
| Balkan                   | 0.43           | 0.04 | 0.62            | 0.15 |
| Caucasus                 | 0.55           | 0.16 | 0.49            | 0.17 |
| E. Transcaucasus         | 0.51           | 0.15 | 0.52            | 0.15 |
| W. steppe periphery      | 0.49           | 0.10 | 0.47            | 0.14 |
| BMAC periphery           | 0.52           | 0.22 | 0.62            | 0.09 |
| E. steppe periphery      | 0.68           | 0.16 | 0.54            | 0.13 |
| N. Australia             | 0.56           | 0.09 | 0.67            | 0.21 |
| W. N America             | 0.65           | 0.13 | 0.64            | 0.19 |
| Mesoamerica              | 0.54           | 0.16 | 0.45            | 0.18 |
| Amazonia                 | 0.58           | 0.13 | 0.58            | 0.20 |
| Mean ± 1 s.d.            | 0.63           | 0.09 | 0.62            | 0.13 |
| Avar sphere              | 0.66           | 0.16 | 0.56            | 0.19 |
| Other N-D                | 0.47           | 0.09 | 0.48            | 0.14 |
| E. steppe near periphery | 0.80           | 0.07 | 0.57            | 0.13 |
| E. steppe far periphery  | 0.57           | 0.15 | 0.52            | 0.14 |
| N. Am. west coast        | 0.68           | 0.14 | 0.68            | 0.22 |
| Intermontane             | 0.62           | 0.14 | 0.60            | 0.18 |

## Findings: High consistency in causative alternation

- Hypotheses:
1. Contact favors high consistency.
  2. Expansive contact favors high consistency.

### 1 Confirmed at face value for animate verbs:

E. steppe periphery (especially inner periphery)

W. North America (especially coastal)

Avar sphere

-- areas where the dominant type is causativizing.

So no support for sheer consistency.

### 2 Confirmed for inanimate verbs:

Balkan, BMAC, N. Australia, W. North America

--some contact zones, not all of them expansive.

Conclusion: Contact can sometimes promote cross-lexeme consistency in derivational paradigms.

## Findings: High frequency of light verb constructions

|                          | % LVC | Std. dev. |
|--------------------------|-------|-----------|
| W. Africa                | 0.03  | 0.06      |
| Balkan                   | 0.00  | 0.00      |
| Caucasus                 | 0.16  | 0.18      |
| E. Transcaucasus         | 0.15  | 0.27      |
| W. steppe periphery      | 0.00  | 0.00      |
| BMAC periphery           | 0.05  | 0.08      |
| E. steppe periphery      | 0.01  | 0.02      |
| N. Australia             | 0.10  | 0.23      |
| W. N America             | 0.00  | 0.02      |
| Mesoamerica              | 0.01  | 0.02      |
| Amazonia                 | 0.00  | 0.00      |
|                          |       |           |
| Mean + 1 s.d.            | 0.12  | 0.18      |
|                          |       |           |
| Avar sphere              | 0.05  | 0.05      |
| Other N-D                | 0.30  | 0.18      |
|                          |       |           |
| E. steppe near periphery | 0.01  | 0.03      |
| E. steppe far periphery  | 0.01  | 0.02      |
|                          |       |           |
| N. Am. west coast        | 0.00  | 0.00      |
| Intermontane             | 0.01  | 0.02      |

Findings: High frequency of light verb constructions

Hypotheses: Contact favors LVC's.  
Expansive contact favors LVC's.

Confirmation strong but partial for both.

What is really going on:

(1) LVC's are common in southwestern Eurasia (maybe southern Eurasia generally).

Also found in Basque, most Iranian, Indo-Aryan.

And northern Australia, and some cluster in Africa (e.g. Hausa).

(2) This wordlist is not ideal for tracking LVC's.

(Better list and survey by Stilo, unpublished? See Stilo 2008.)

Higher frequencies of LVC's are accompanied by higher standard deviations. This shows that high LVC frequency is not very consistent even where it is an areal feature, suggesting that as a typological variable it is recessive overall.

## Findings: Noun-based derivational morphology

Survey is very labor-intensive

(200-item wordlist, entry and derivational history for each)

Therefore, start with close internal survey of one area:

Avar sphere (North central-eastern Caucasus)

Geographical conditions turn out to be relevant:

Northern and southern slopes

Northern and southern economic orientations

defined by winter pastures

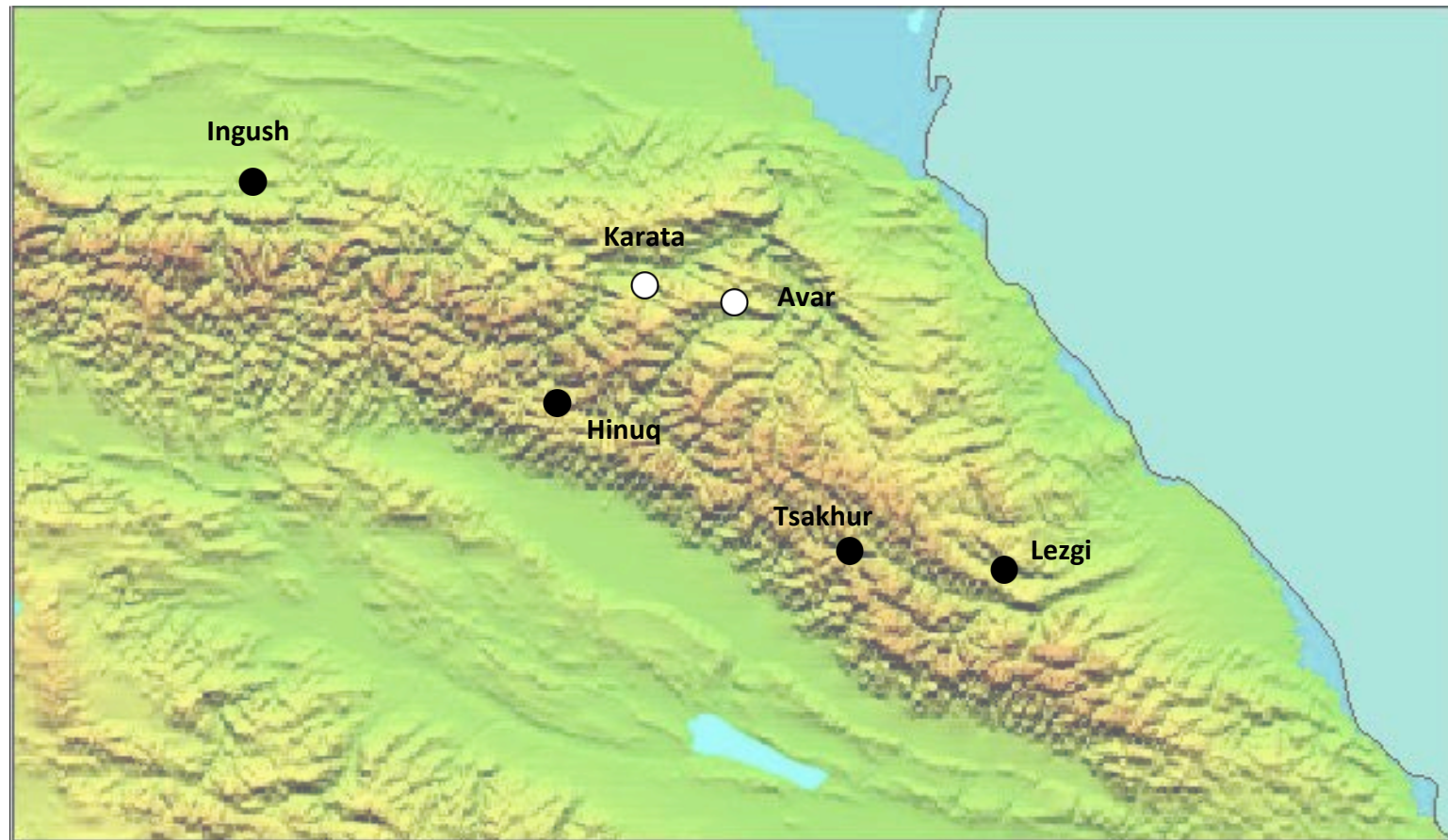
N = toward steppe

S = toward Transcaucasus

## Findings: Noun-based derivational morphology

|               | % N-based | Slope | Econ. |
|---------------|-----------|-------|-------|
| Ingush        | 0.82      | N     | N/S   |
| Avar          | 0.62      | N     | N     |
| Karata        | 0.65      | N     | N     |
| Hinuq         | 0.79      | N     | N/S   |
| Tsakhur       | 0.81      | N/S   | S     |
| Lezgi         | 0.83      | N/S   | N/S   |
|               |           |       |       |
| cf. Kabardian | 0.33      | N     | N     |

### Percent of items that are noun-based



○ Other    ○ Lower    ● High

## Findings: Noun-based derivational morphology

Hypotheses: Noun-based is favored in contact situations.

Noun-based is favored in expansive contact situations.

Outcome:

% noun-based is higher in southern-oriented economies.

(Part of same Southern Eurasian LVC-using area.)

Strong family signatures:

Nakh-Daghestanian is strongly noun-based

West Caucasian is strongly verb-based and flexible



## Findings: High POS flexibility of derivational base

|                            | % flexible<br>(N=V) |
|----------------------------|---------------------|
| Ingush                     | 0                   |
| Avar                       | 0                   |
| Karata                     | 0                   |
| Hinuq                      | 0.03                |
| Tsakhur                    | 0                   |
| Lezgi                      | 0                   |
|                            |                     |
| cf. Kabardian              | 0.36                |
|                            |                     |
| cf. Tungusic (E. Siberia): |                     |
| Evenki                     | 0.36                |
| Nanai                      | 0.03                |

Findings: High POS flexibility of derivational base

Hypotheses:

(Foley) Flexibility is selected against in contact.

(alternative) Flexibility is favored in isolation.

Outcome:

Nakh-Daghestanian family profile: Near-zero flexibility.

Not affected by contact, north vs. south orientation, etc.

E. Siberia:

Generally, flexibility is low in Mongolia and nearby;

higher in far north and east

Tungusic: low in Nanai (settled, in Mongolic sphere);

high in Evenki (nomadic reindeer herders)

Interpretation:

Flexibility is selected **against** in contact (confirming Foley)

## Summary: Derivational type and contact situation

### *Factor*

Causativization

Consistency of caus. alt.

Light verb constructions

Noun-based derivation

POS flexibility

### *Favored by*

Expansive contact

Contact (any);

    Specific type (causativizing)

Specific contact (Transcaucasus)

Specific contact (Transcaucasus)

    Inheritance (sets limits)

**Disfavored by** contact

    Inheritance (sets limits)

## Conclusions

*Derivational type (base transitivity, noun vs. verb base):*

Genealogically fairly stable: Strongly inherited, limits extent of response to contact

*Implication for hybridization:*

Where derivational type or its symptoms diffuse areally, genealogical and typological relatedness are hard to untangle. This happens especially under frontier conditions. Clear case: eastern steppe, where genealogy and typology among Turkic, Tungusic, and Mongolic have taken long to untangle.

*Survey:*

Labor-intensive but worthwhile: Can discriminate between

- heredity and contact
- expansive vs. other contact
- contact per se and effects of specific type in contact

## Further work

### Fine-tune wordlist

- Also, develop a short wordlist for quick-and-dirty assessments

- Fine-tune sociolinguistic typology and pigeonholing of languages

### Expand language survey:

- Intensive: Caucasus, eastern steppe, western North America

- Extensive: everywhere

### Expand the typological variables (including measures of complexity)

### Analysis and theory:

- Factor out effects of sociolinguistics, inherent genealogical stability of variables, rates of diffusion, rates of change

- Improve typological analysis of verb-adjective flexibility

- Cross-linguistic survey of inherent stability of variables

Pie in sky: Change the way descriptive linguistics presents lexical data!

## References

- Nedjalkov, V. P. 1969. Nekotorye verojatnostnye universalii v glagol'nom slovoobrazovanii. In F. Vardul', ed., *Jazykovye universalii i lingvističeskaja tipologija*, 106-114. Moscow: Nauka.
- Nichols, Johanna. 1982. Ingush transitivity and detransitivization. *BLS* 8.445-462.
- Nichols, Johanna, David A. Peterson, and Jonathan Barnes. 2004. Transitivity and detransitivizing languages. *Linguistic Typology* 8:2.149-211.
- Stilo, Don. 2008. An introduction to the Araxes-Iran linguistic area. Presented at SOAS, December 2008.  
<http://www.soas.ac.uk/linguistics/events/deptseminars/02dec2008-an-introduction-to-the-araxes-iran-linguistic-area.html>
- Thompson, Laurence C. and M. Terry Thompson. 1996. *Thompson River Salish Dictionary*. Missoula: University of Montana.