Visualizing Valency Alternations

Michael Cysouw[1] & Iren Hartmann[2]

[1]Universität Marburg, [2]Max-Planck-Institut für evolutionäre Anthropologie

DGfS Tagung 2013, Potsdam

13. 3. 2013

Outline

i. Valency Alternations
ii. Data
iii. Visualizations
iv. Summary

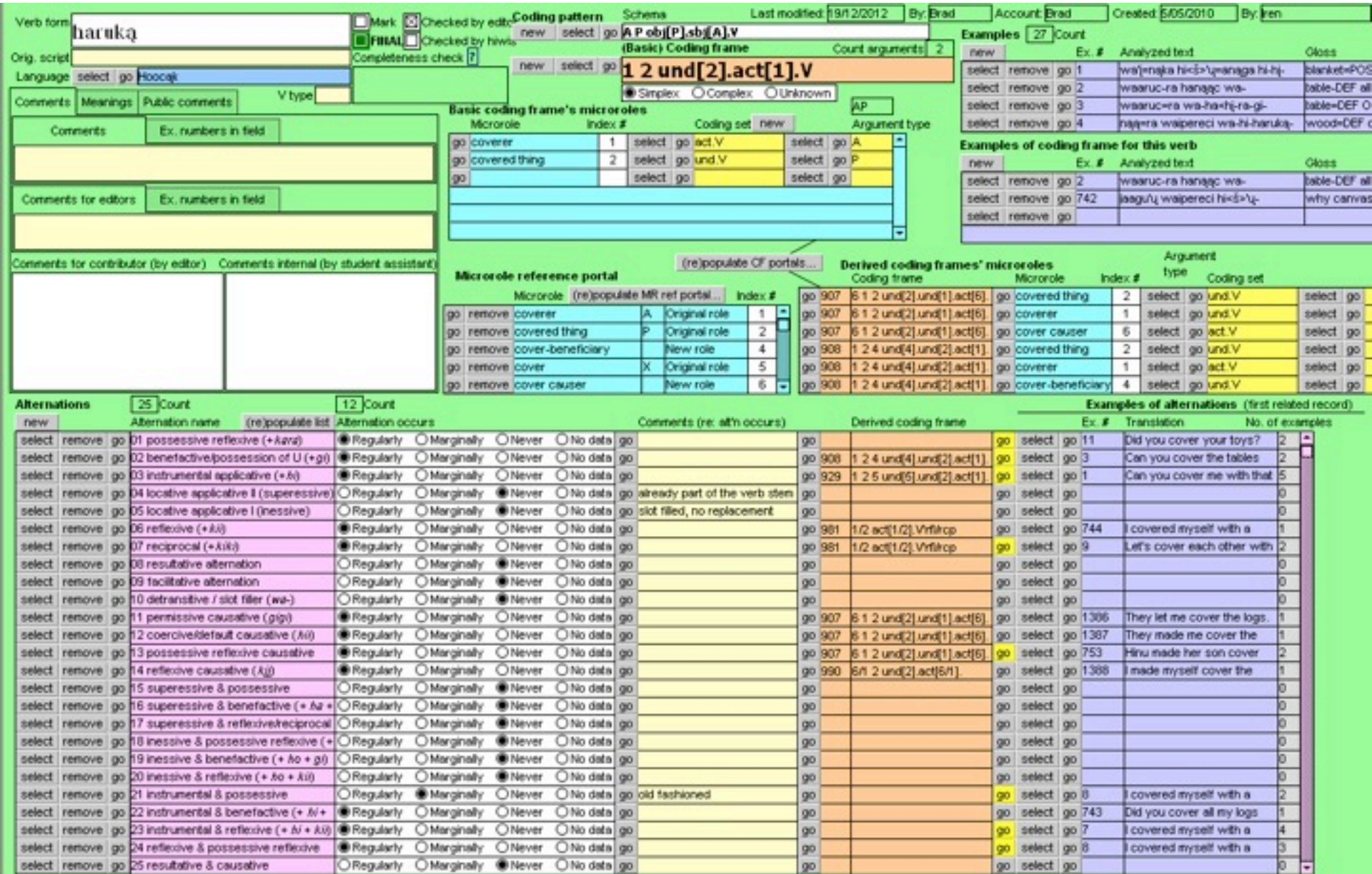
Valency Alternations

- valency alternations have been **studied** and described comprehensively **for** some **individual languages**
- **no** real attempt has been made to perform a **large-scale cross-linguistic study**
- our current project collects comparable data for valency patterns and alternations in **30 diverse languages worldwide**
- in our presentation today, detailed data on alternations from 11 languages is used

Project languages (for basemap)

L's with more detailed alternation data

Data collection



Data

- counterparts ('translations') from 35 languages for 70 **verb meanings** (conceived of as representative of the verbal lexicon)
- one primary valency frame (**coding frame**) and all possible alternative valency frames (**derived coding frames**) for each counterpart
- coding frames include information on argument coding (flags & indexes)
- arguments are assigned to **microroles** ('verb-specific roles') and **coding sets**

Example from German

Meaning Label: EAT

Verb form: *essen*

Coding frame: 1-nom V.sbj[1] 2-acc

Microroles:	EATER(1)	EATEN THING(2)
Coding Sets:	nom & V.sbj	acc

Example: *Das Kind isst den Apfel.*

Alternation 1: Passive: ✓ *(Der Apfel wurde (von dem Kind) gegessen.)*

Derived Frame: 2-nom passV‘.sbj[2]

Coding Set: nom & V.sbj

Alternation 2: Object Omission: ✓ *(Das Kind isst.)*

Derived Frame: 1-nom V‘.sbj[1]

Coding Set: nom & V.sbj

Alternation 3: *be-* Applicative: ✘ *(*Das Kind be-isst den Apfel.)*

Alternations and Alternants

Alternation → relation between two alternants

- e.g. Causative:

1 V	(*1 runs*) →	3 1 V‘	(*3 causes 1 to run*)
1 V 2	(*1 breaks 2*) →	3 1 V‘ 2	(*3 causes 1 to break 2*)

- Alternants (a.k.a Diatheses or Voices)
- Alternants can consist more than one coding frame (a.k.a. subcategorization frame)

Comparative and L specific concepts

Comparative Concepts:

- microroles (EATER, HITTER, LIKER, RUNNER, …)
- verb meanings (EAT, HIT, LIKE, RUN, …)

Language Specific Concepts:

- coding frames (1-nom 2-acc V.sbj[1], …)
- coding set (nom, acc, sbj.V, obj.V, *við*&acc…)
- alternant (“voice”) (set of coding frames)

Visualization

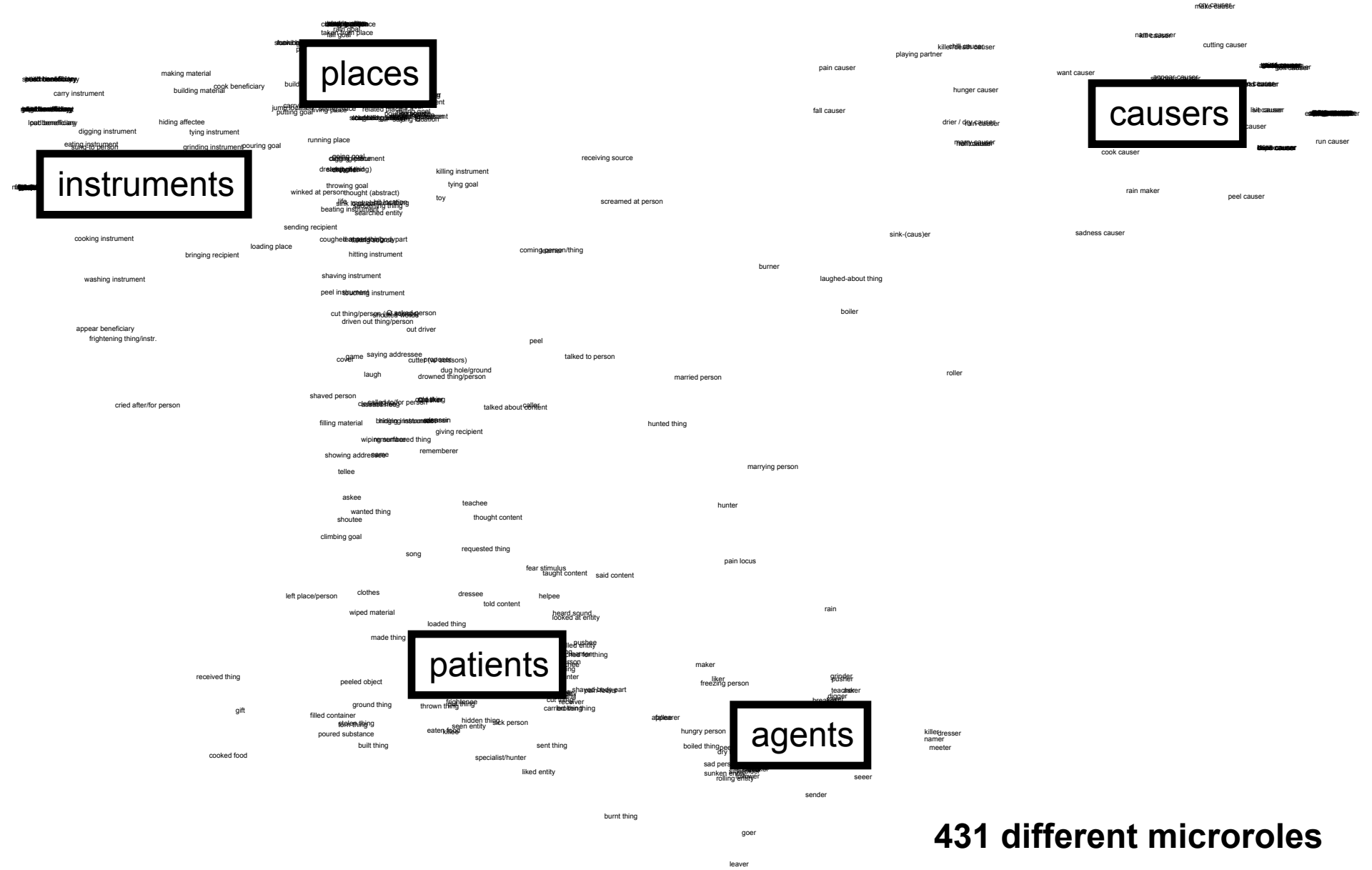
- How can we visualize linguistic diversity?
  - Show differences/similarities between languages
  - Basic idea: use a **constant "universal" base map**, and show each language as a separate overlay
  - **Optimize base map to highlight variation** among the languages to be compared
  - With sufficient data, suitable base maps approach an average ("universal") language structure
  - Construction of a base map is similar to the semantic map approach from linguistic typology

Visualization

- Three aspects to be visualized:
  - Visualize **alternants** (“voices”) individually
  - Visualize **alternations** (“relations between alternants”)
  - Visualize **similarity between alternations** across all languages (“typology of alternations”)

1. Alternants

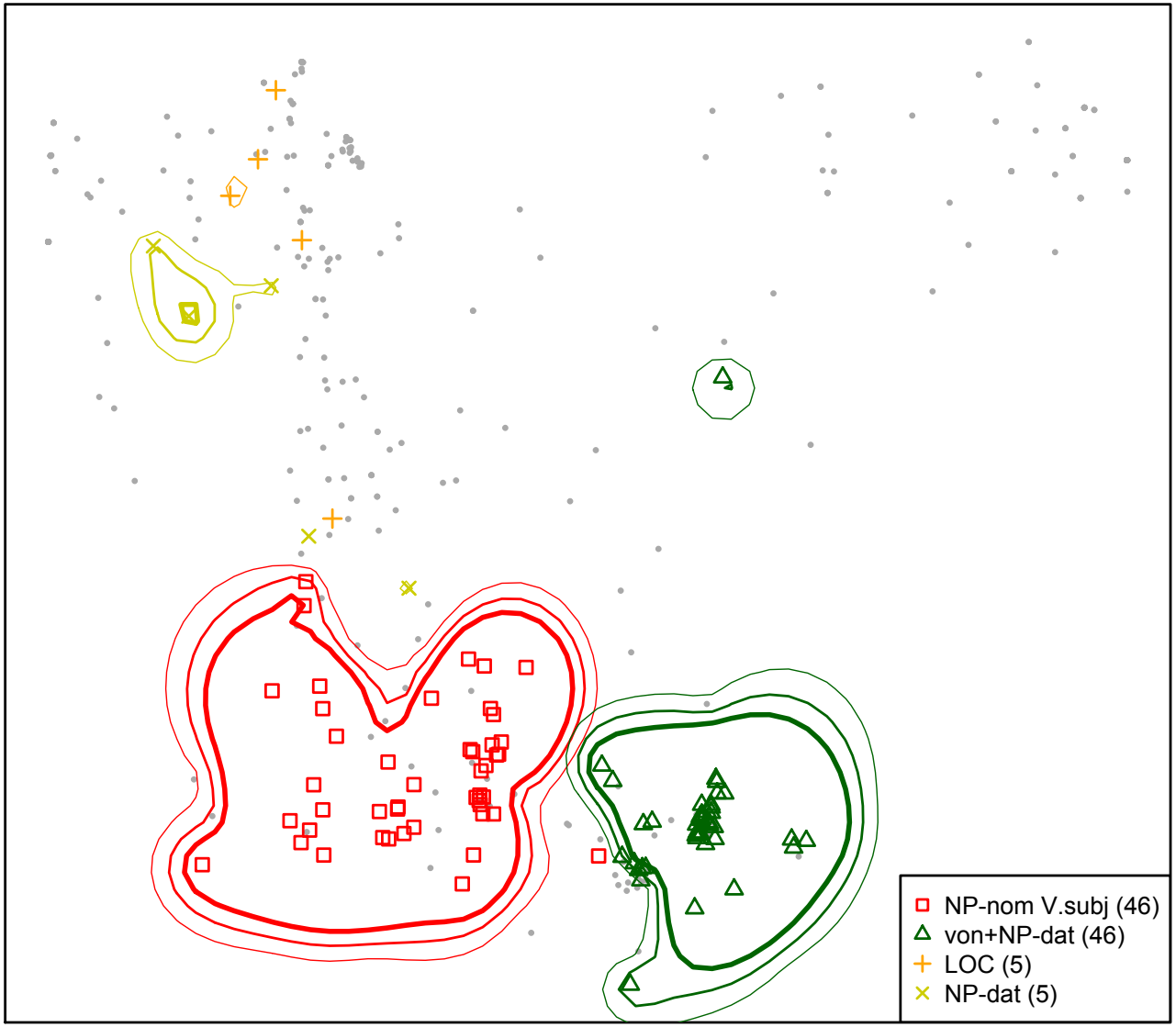
- An **alternant** (e.g. German Passive marked with *werden+Partizip II*) is characterized by a **mapping between microroles** ("lexical roles") **and coding sets** ("language specific marking").
  - HITTER, THROWER, etc ("agents") are marked with prepositional phrase *von + NP + Dative case*
  - HIT THING, THROWN THING, etc ("patients") are marked with *Nominative case + Verb inflection*
- For **base map**, use **microroles**, and **position** them **according to average co-expression**



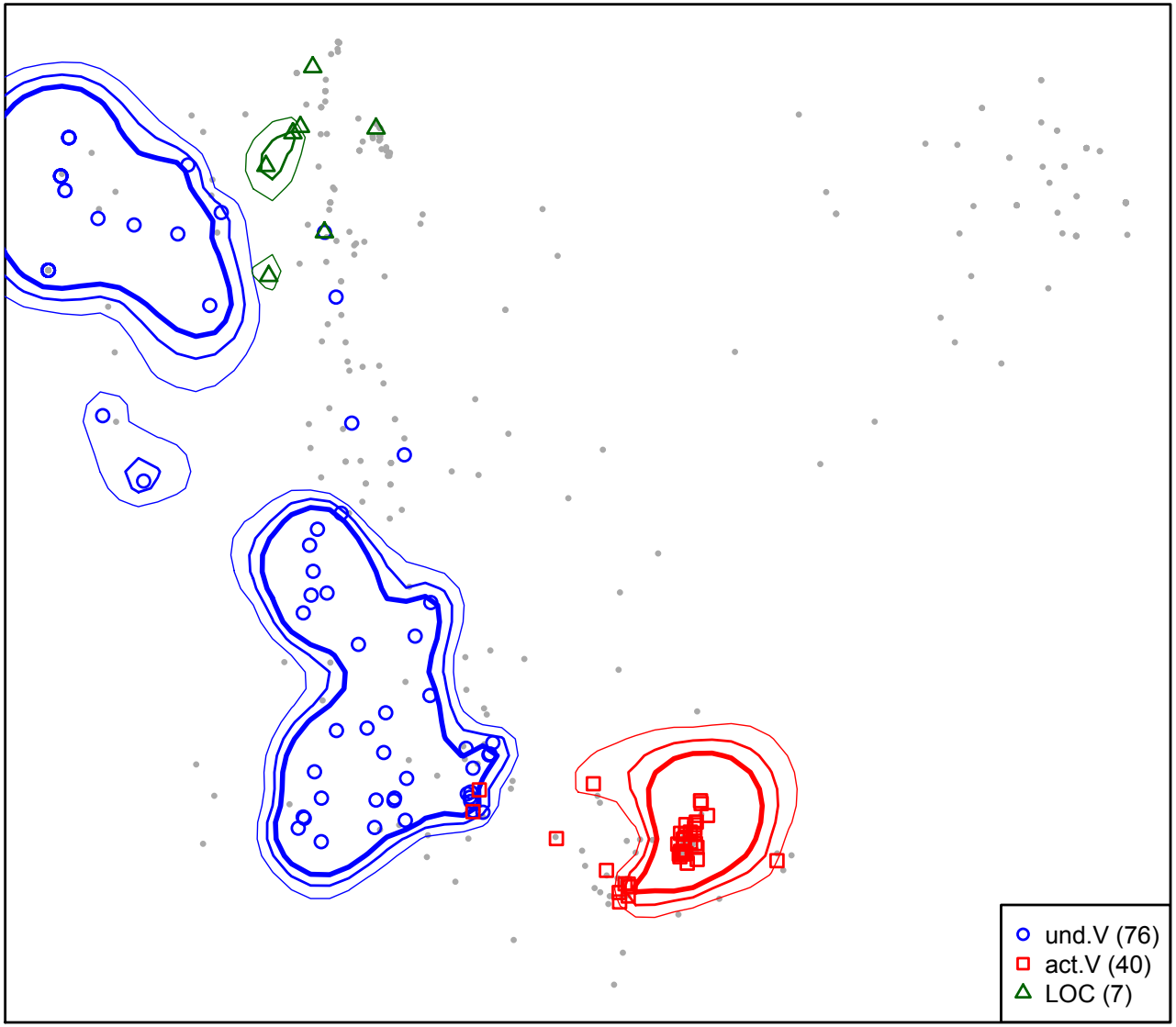
431 different microroles

Position determined by first two dimensions of singular vector decomposition of normalized microrole similarity measured by multilingual coexpression

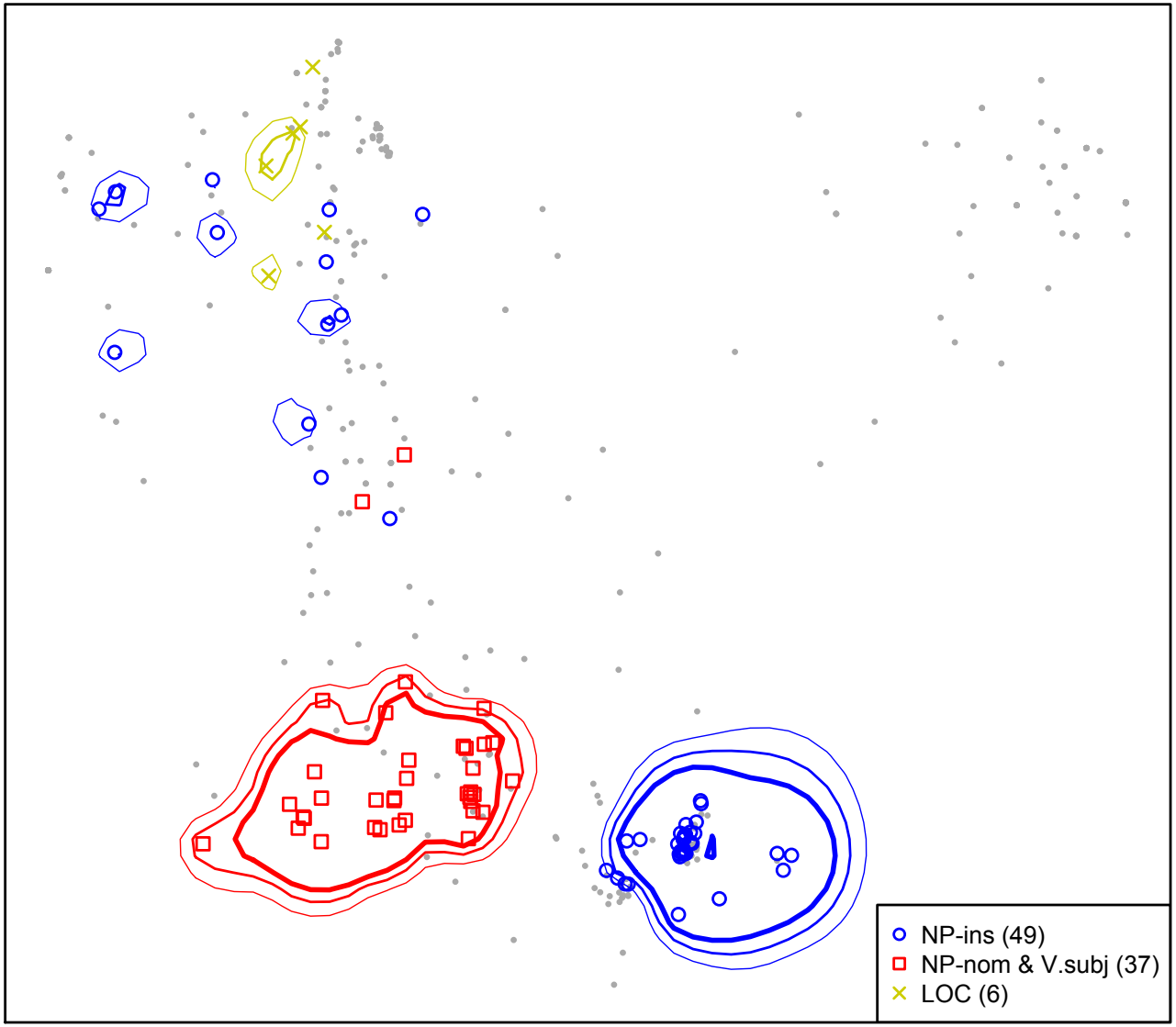
Passive (werden)_German



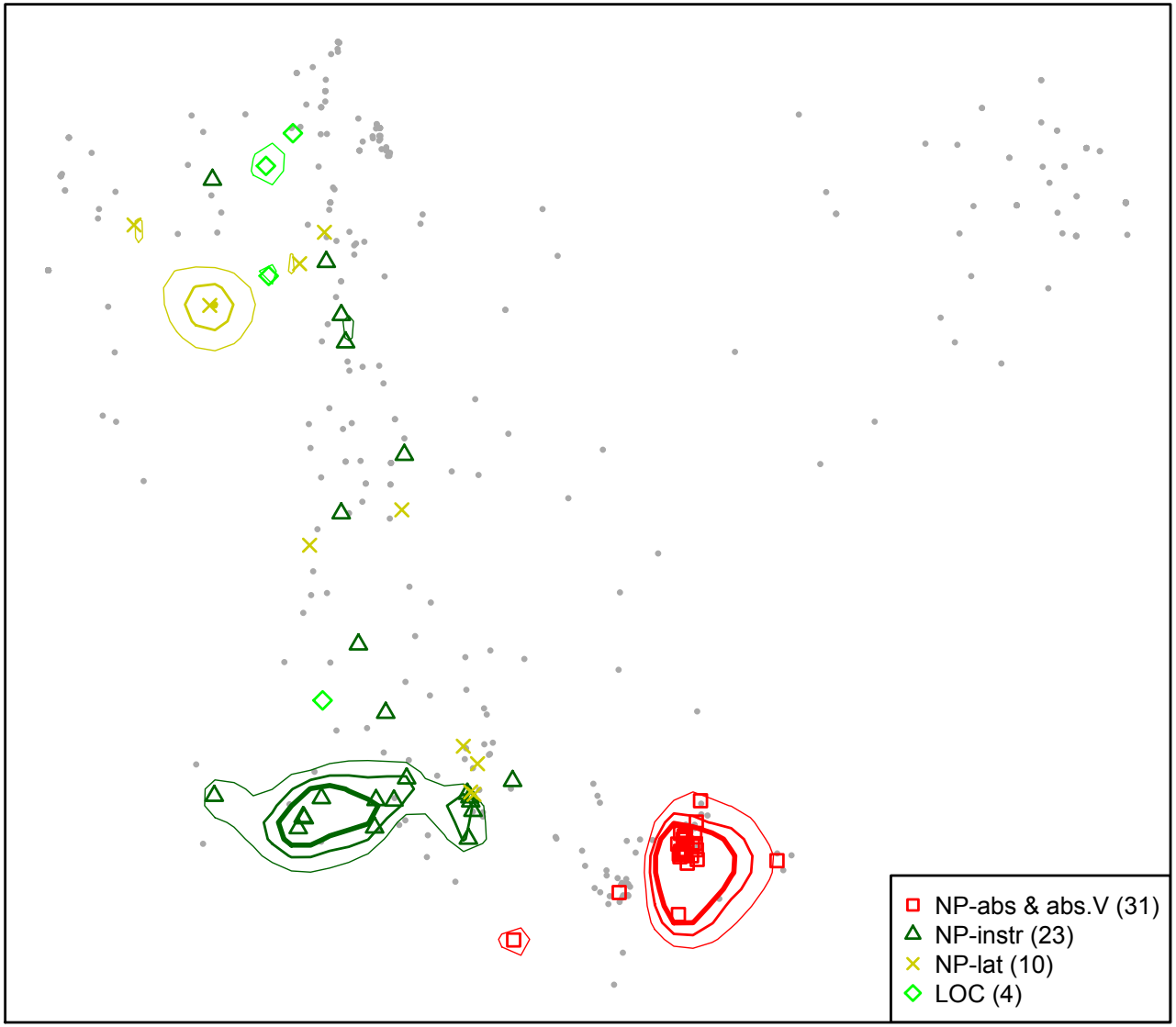
02 benefactive/possession of U (+gi)_Hoocąk



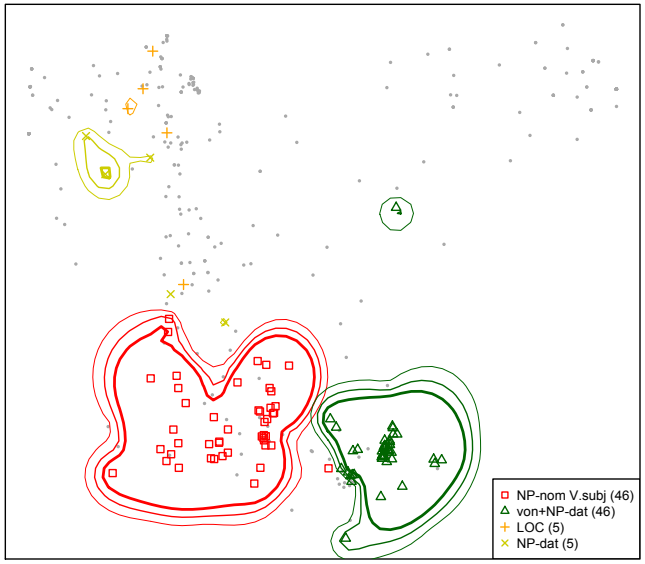
Participial Passive_Russian



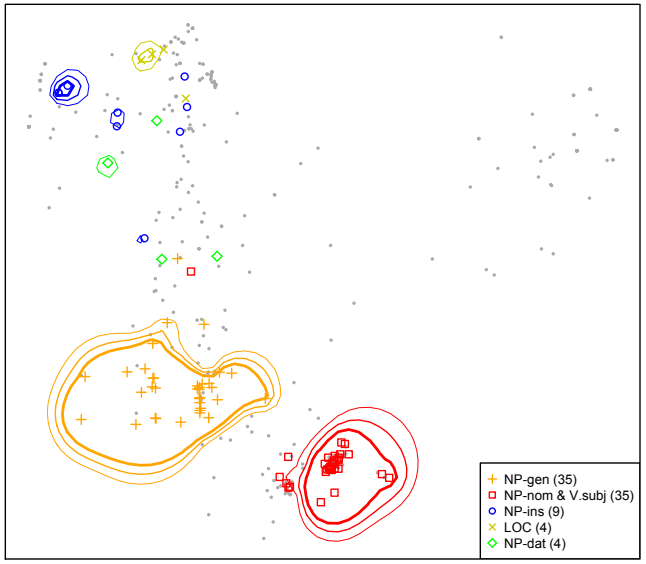
Antipassive Alternation_Bezhta



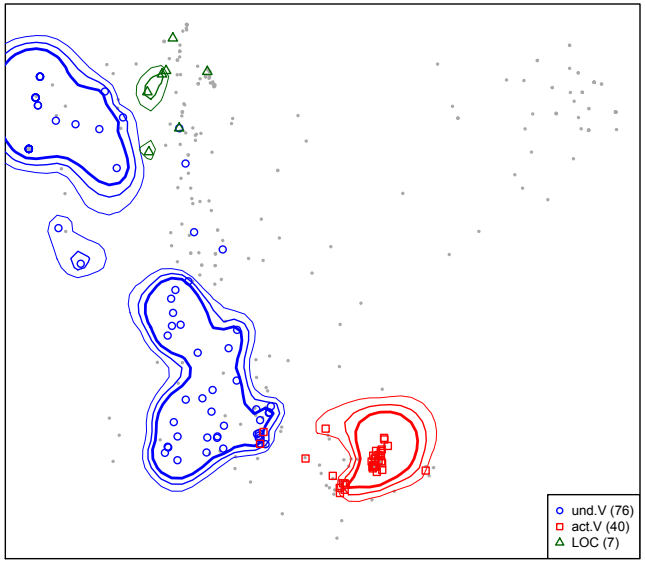
Passive (werden)_German



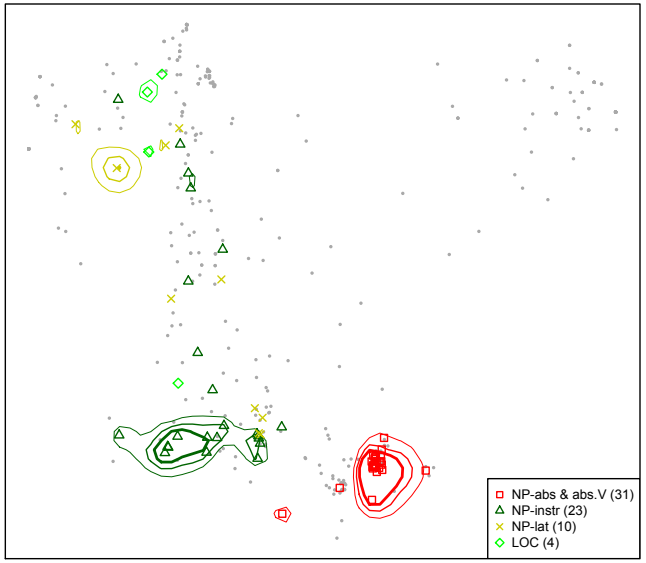
Negative Accusative-Genitive alternation_Russian



02 benefactive/possession of U (+gi)_Hoocąk



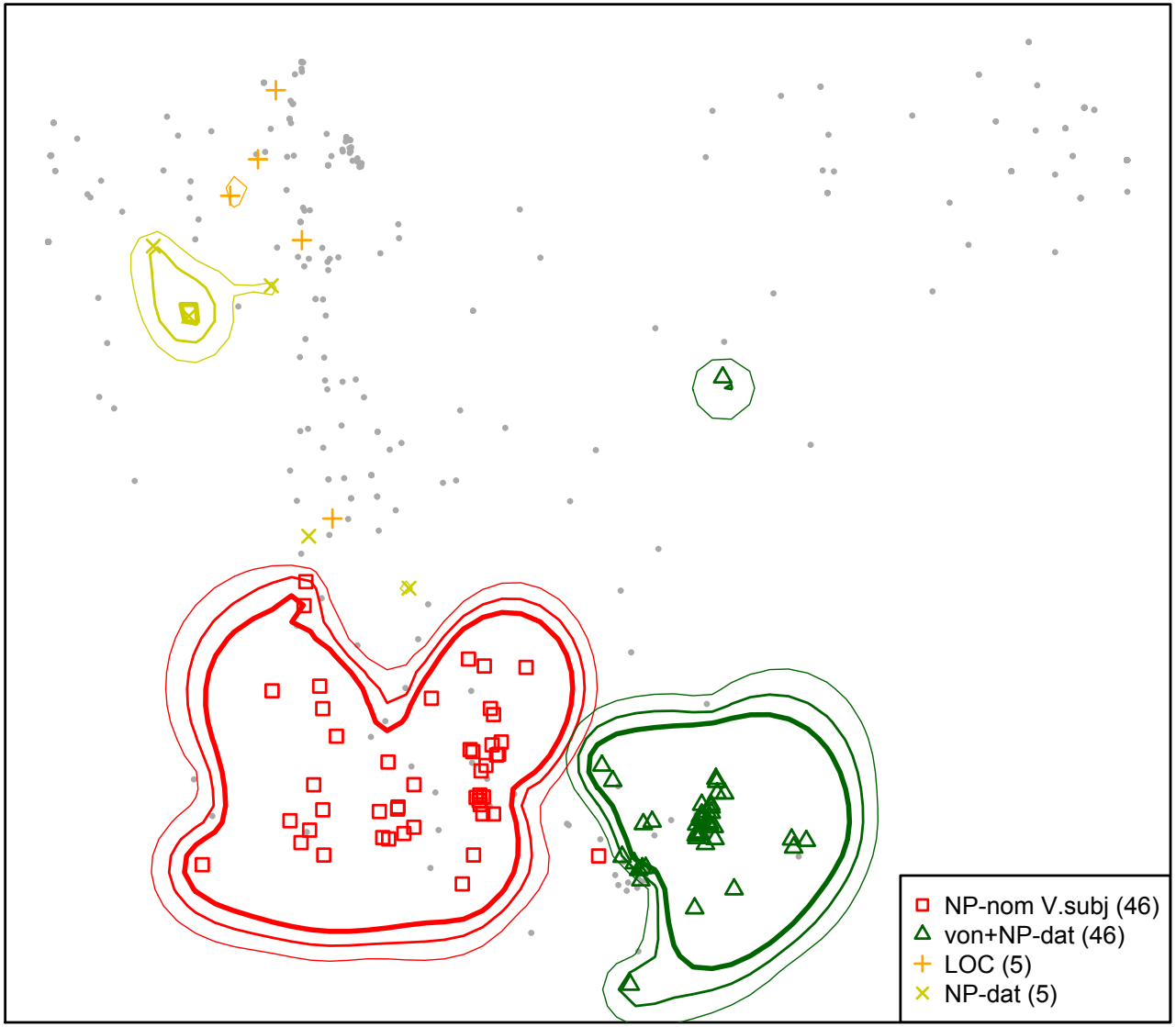
Antipassive Alternation_Bezhta



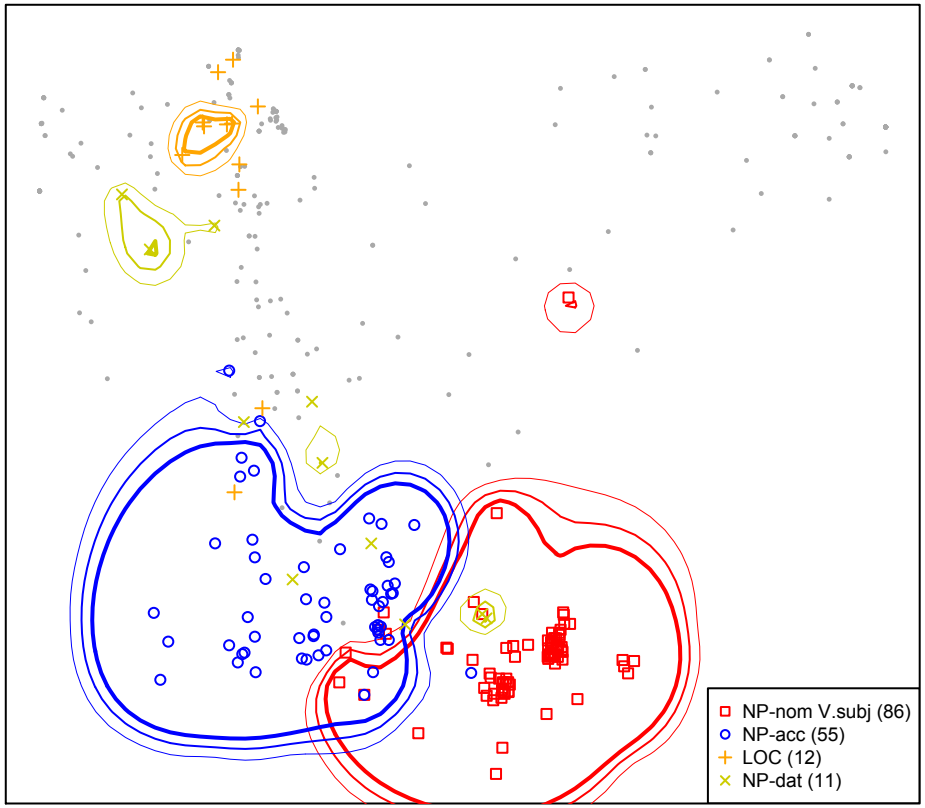
2. Alternations

- These visualizations **cannot be directly compared**
  - clusters of microroles are one important notion, but the **actual coding is also important**
  - but: **how to compare language-specific coding?** e.g. what is the relation between
    - German *von+NP-dat,* Hoocąk *act.V,* Russian *NP-ins,* Bezhta *NP-abs & abs.V*
- Our proposal: **compare alternant to 'unmarked' alternant of the same language**

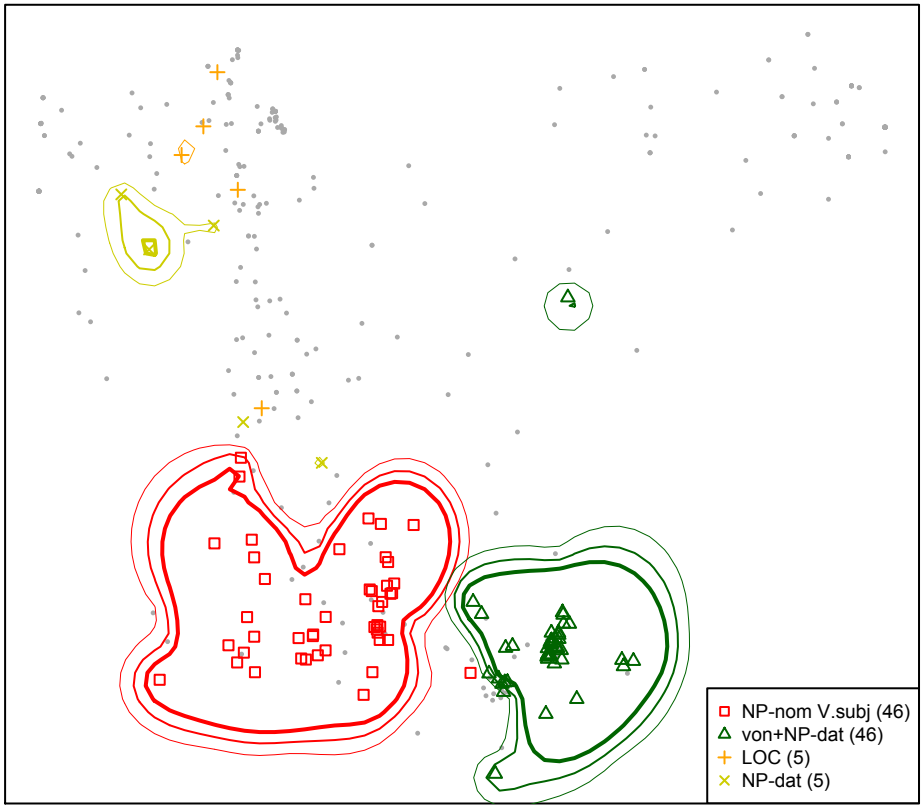
Passive (werden)_German



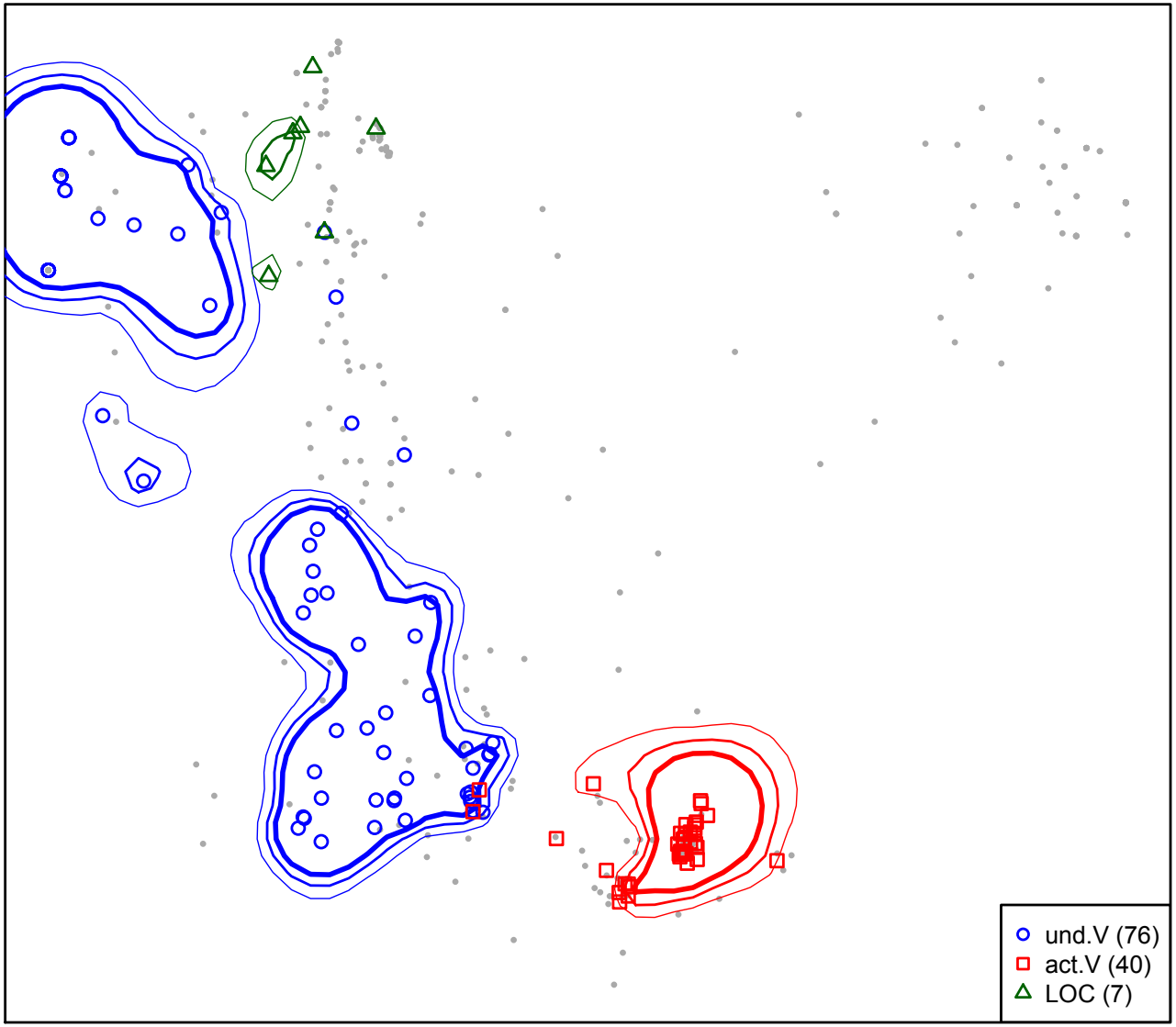
_German



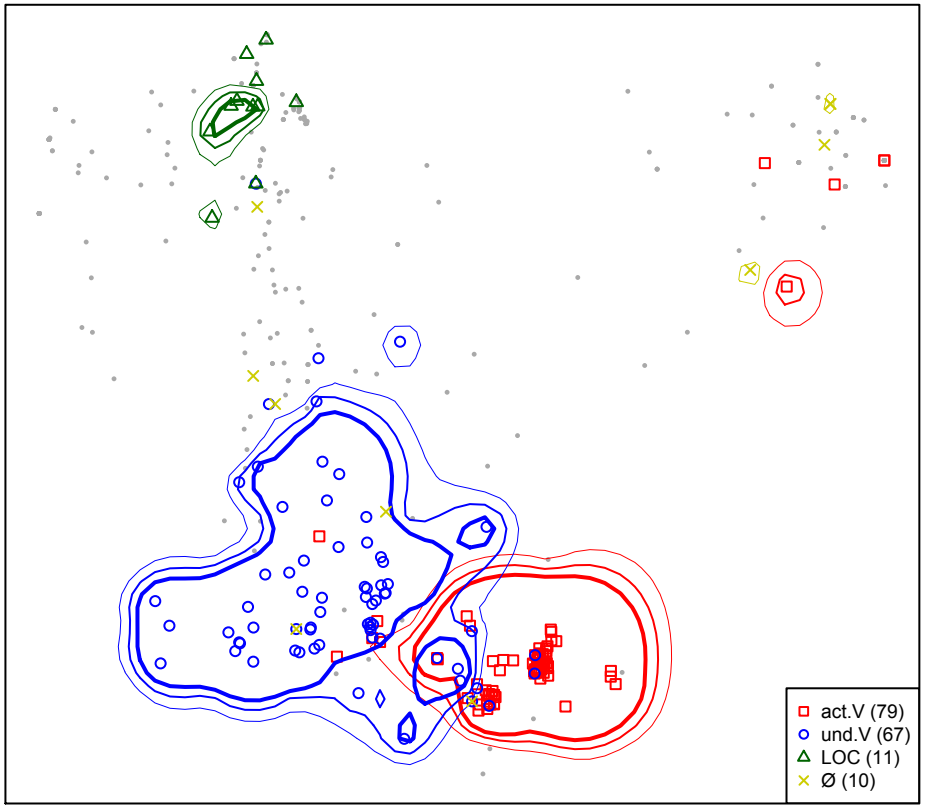
Passive (werden)_German



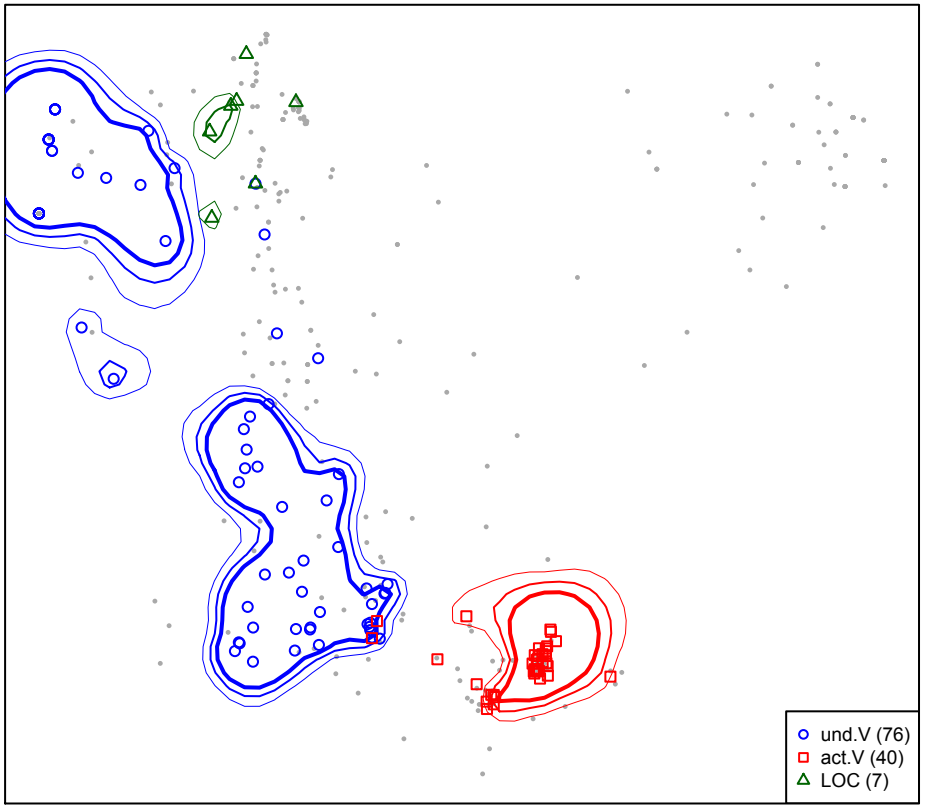
02 benefactive/possession of U (+gi)_Hoocąk



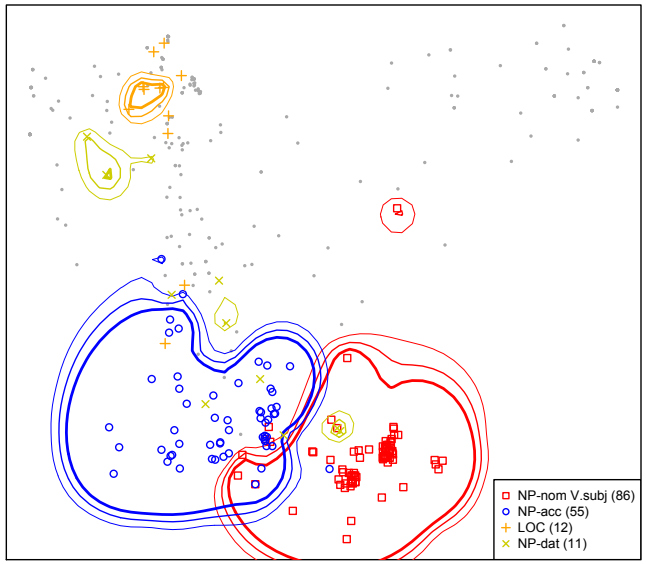
_Hoocąk



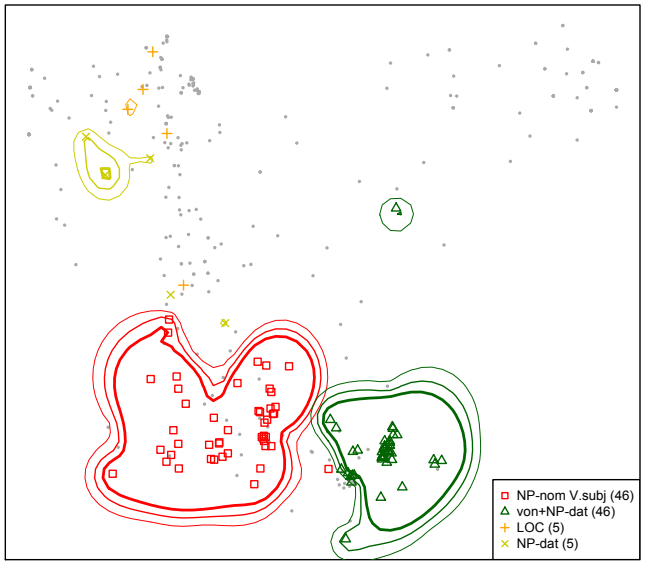
02 benefactive/possession of U (+gi)_Hoocąk



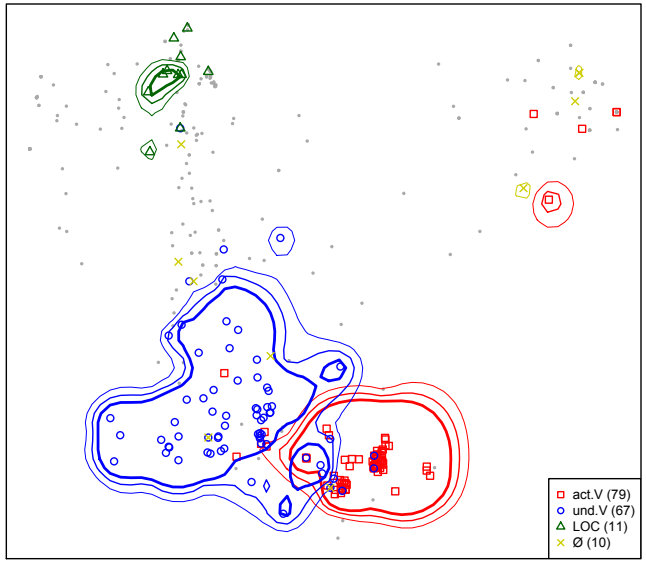
_German



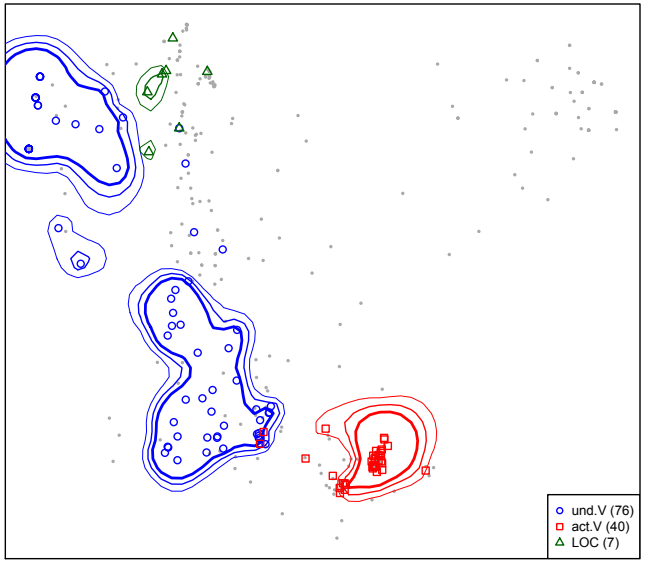
Passive (werden)_German



_Hoocąk



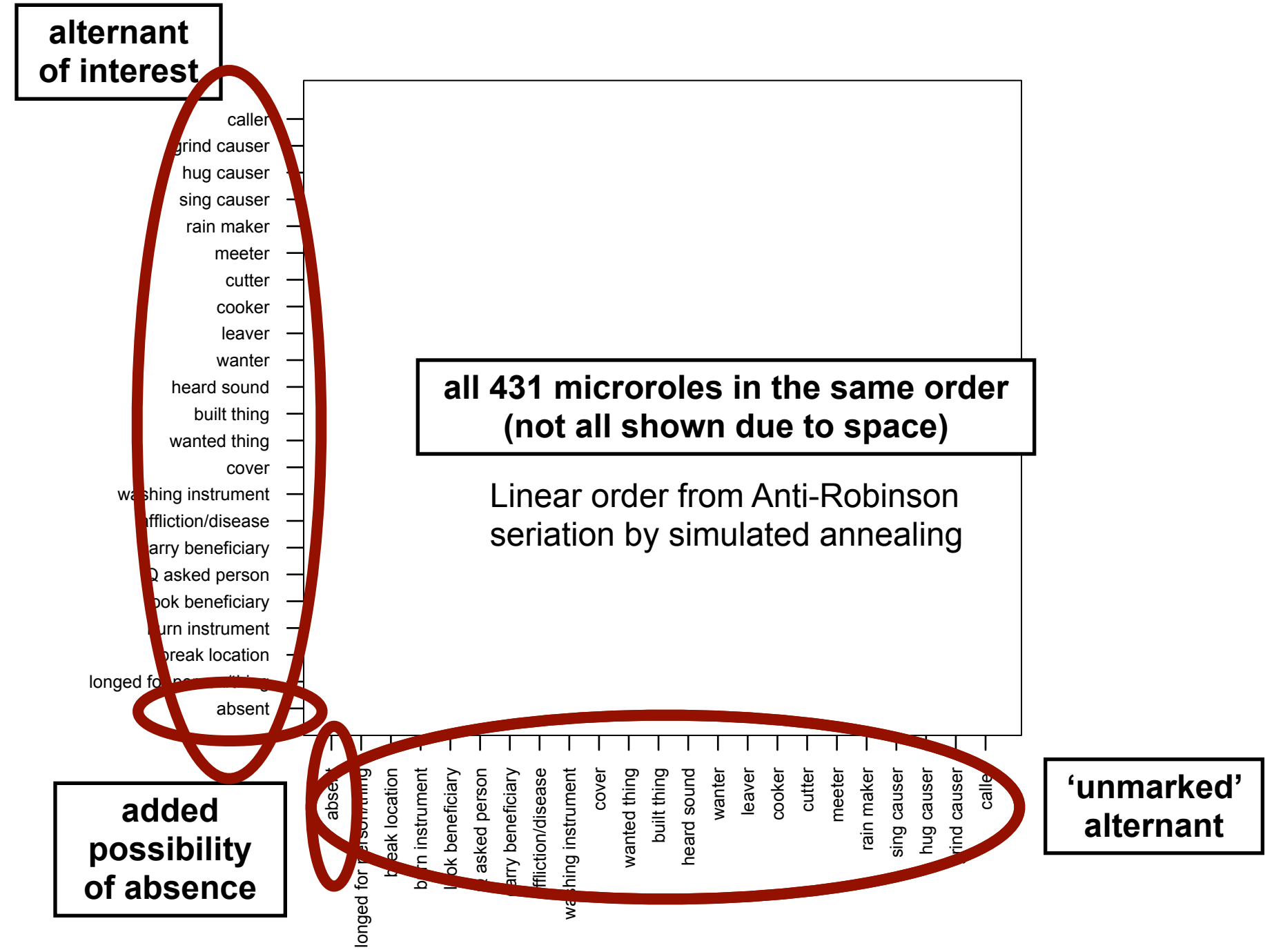
02 benefactive/possession of U (+gi)_Hoocąk



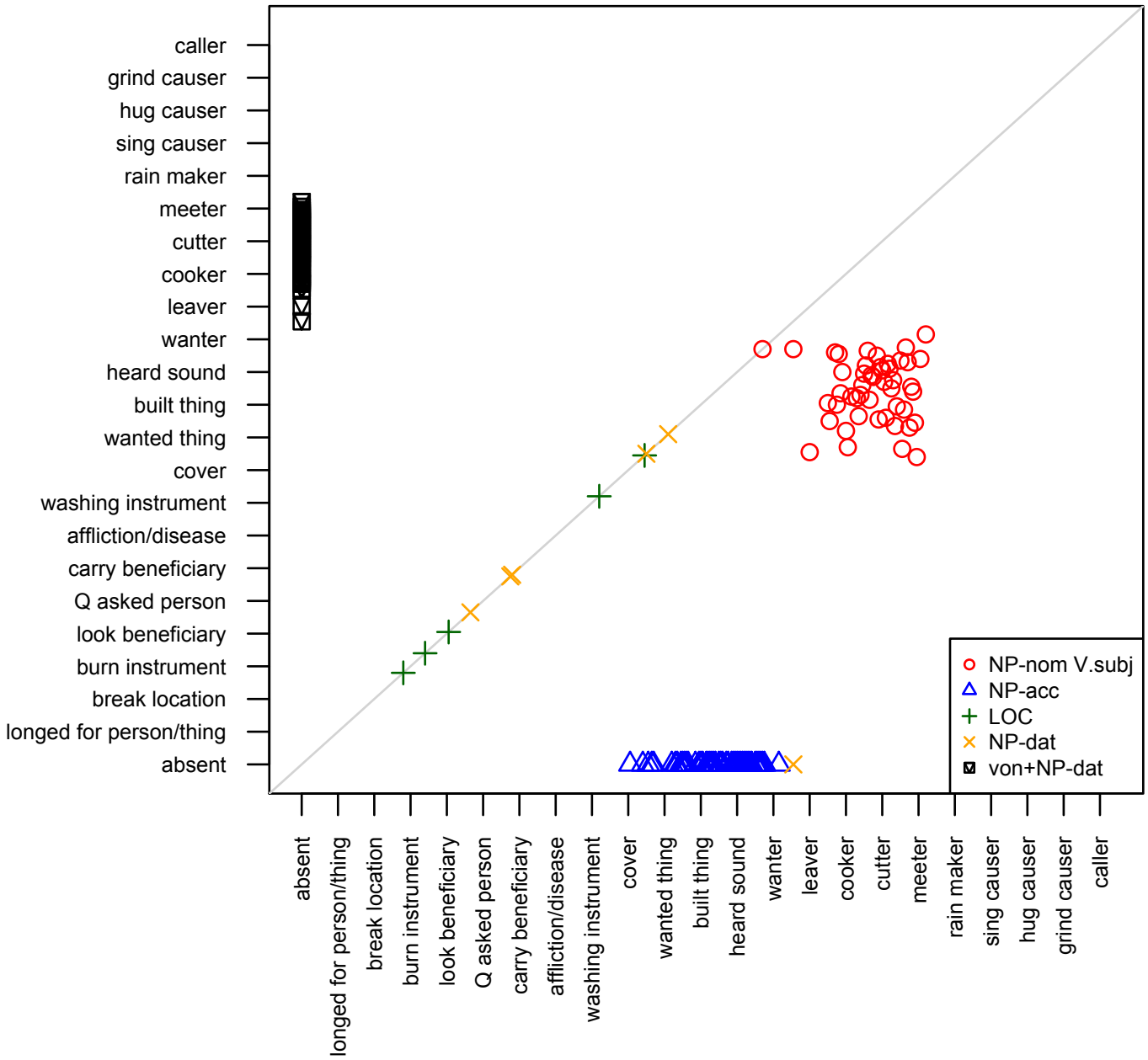
Difficult to interpret ...

Visualize Alternation

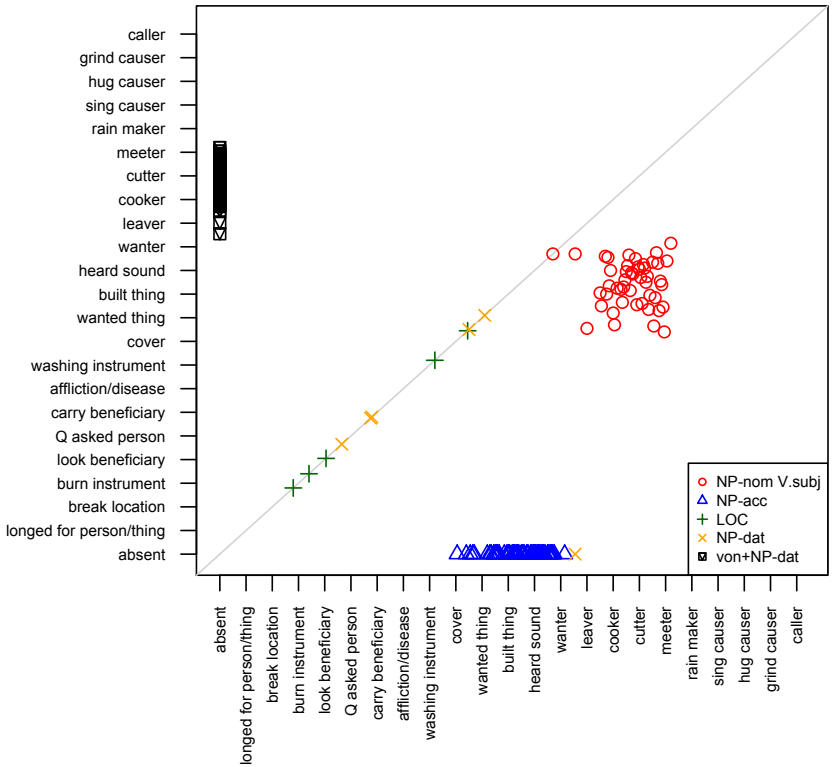
- Another 2-dimensional representation:
  - Use horizontal axis for ‘unmarked’ alternant
  - Use vertical axis for alternant of interest
  - Each axis shows all 431 roles, always in exactly the same order (one-dimensional base map)


alternant of interest
caller
grind causer
hug causer
sing causer
rain maker
meeter
cutter
cooker
leaver
wanter
heard sound
built thing
wanted thing
cover
shing instrument
ffliction/disease
arry beneficiary
Q asked person
ook beneficiary
urn instrument
reak location
absent
all 431 microroles in the same order (not all shown due to space)
Linear order from Anti-Robinson seriation by simulated annealing
absent
cover
wanted thing
built thing
heard sound
wanter
leaver
cooker
cutter
meeter
rain maker
sing causer
hug causer
added possibility of absence
'unmarked' alternant

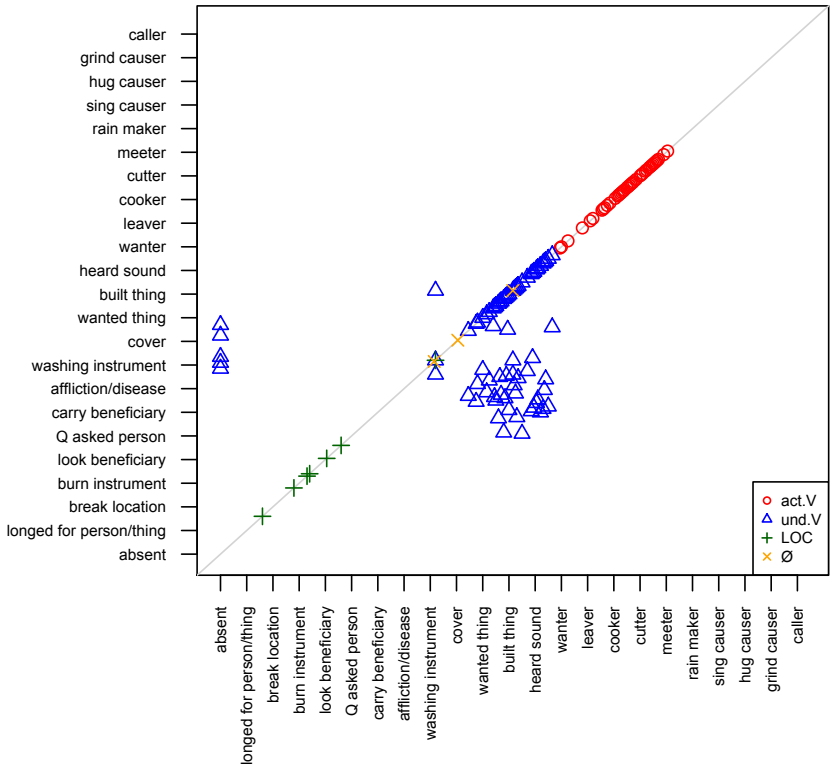
Passive (werden) from German



Passive (werden) from German



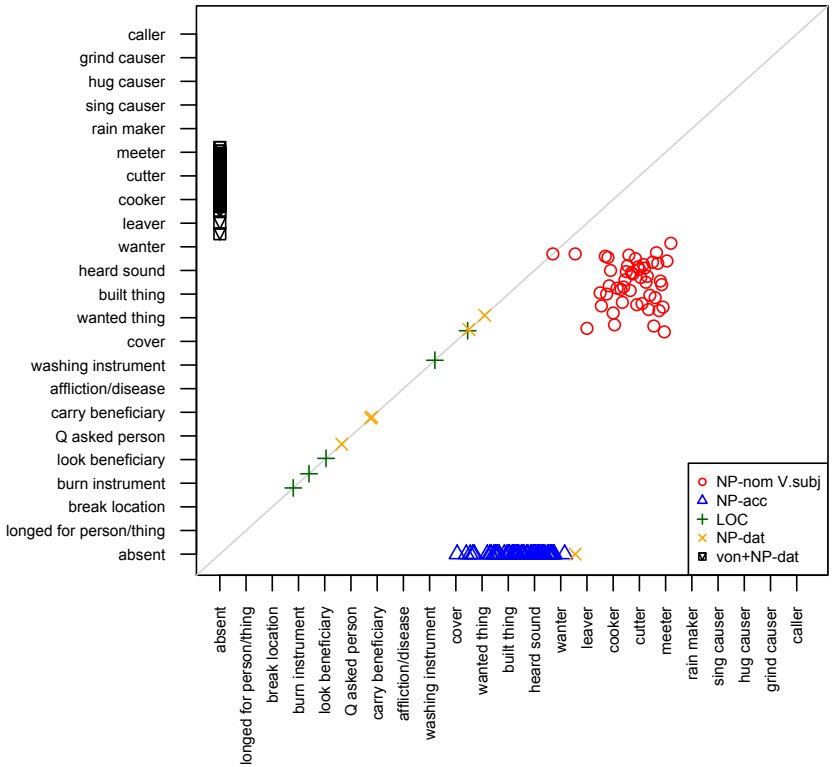
02 benefactive/possession of U (+gi)_Hoocąk



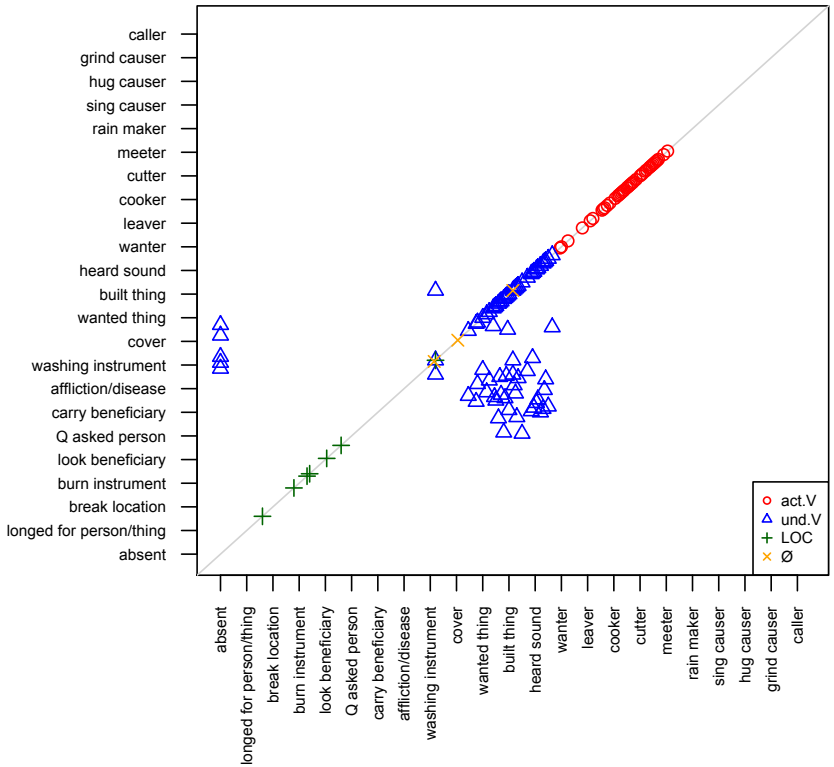
3. Typology of Alternations

- These visual representation of **Alternations can directly be compared**
- Note that the **color is not important** anymore!
- Simple measure of similarity between to alternations is the **number of overlapping dots**
- This similarity can be clustered to show typological structure

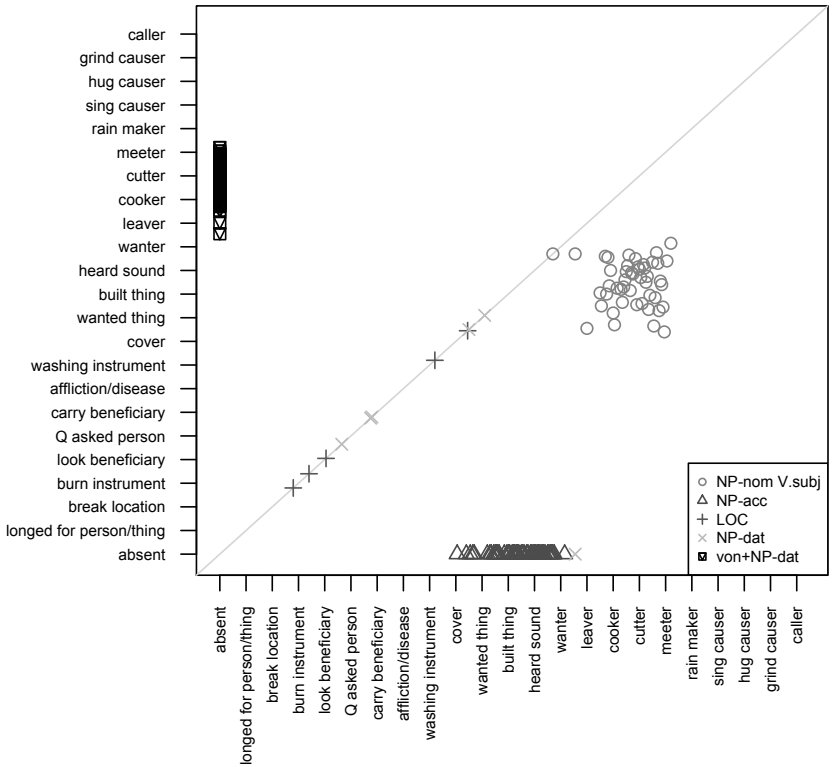
Passive (werden) from German



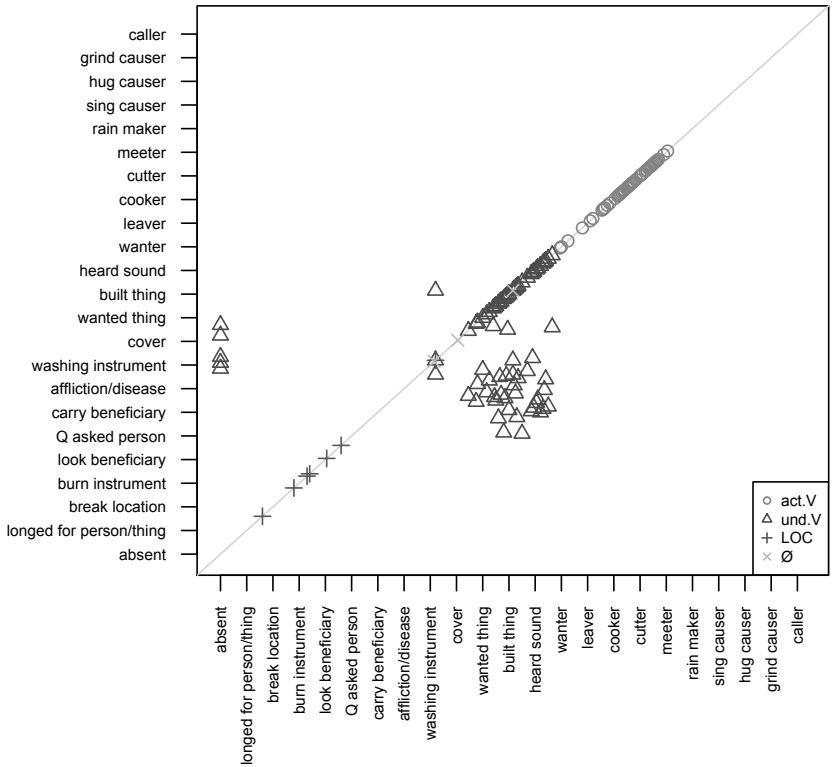
02 benefactive/possession of U (+gi)_Hoocąk



Passive (werden) from German

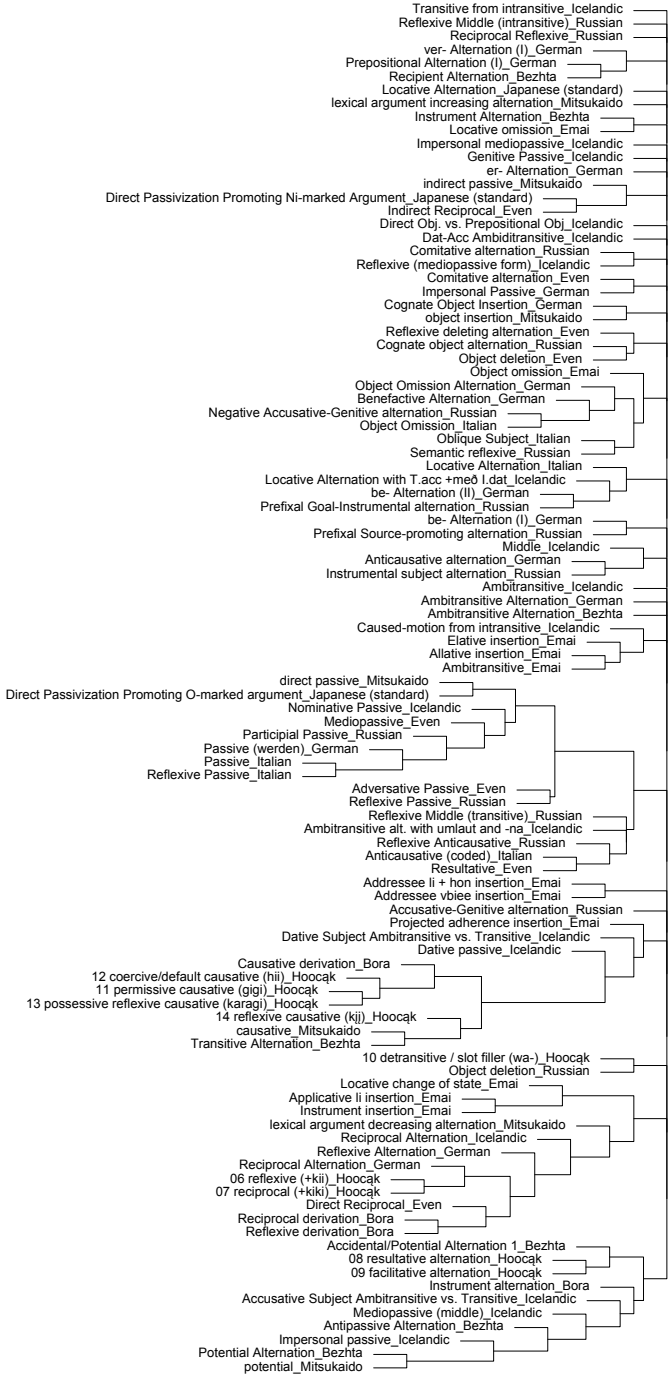


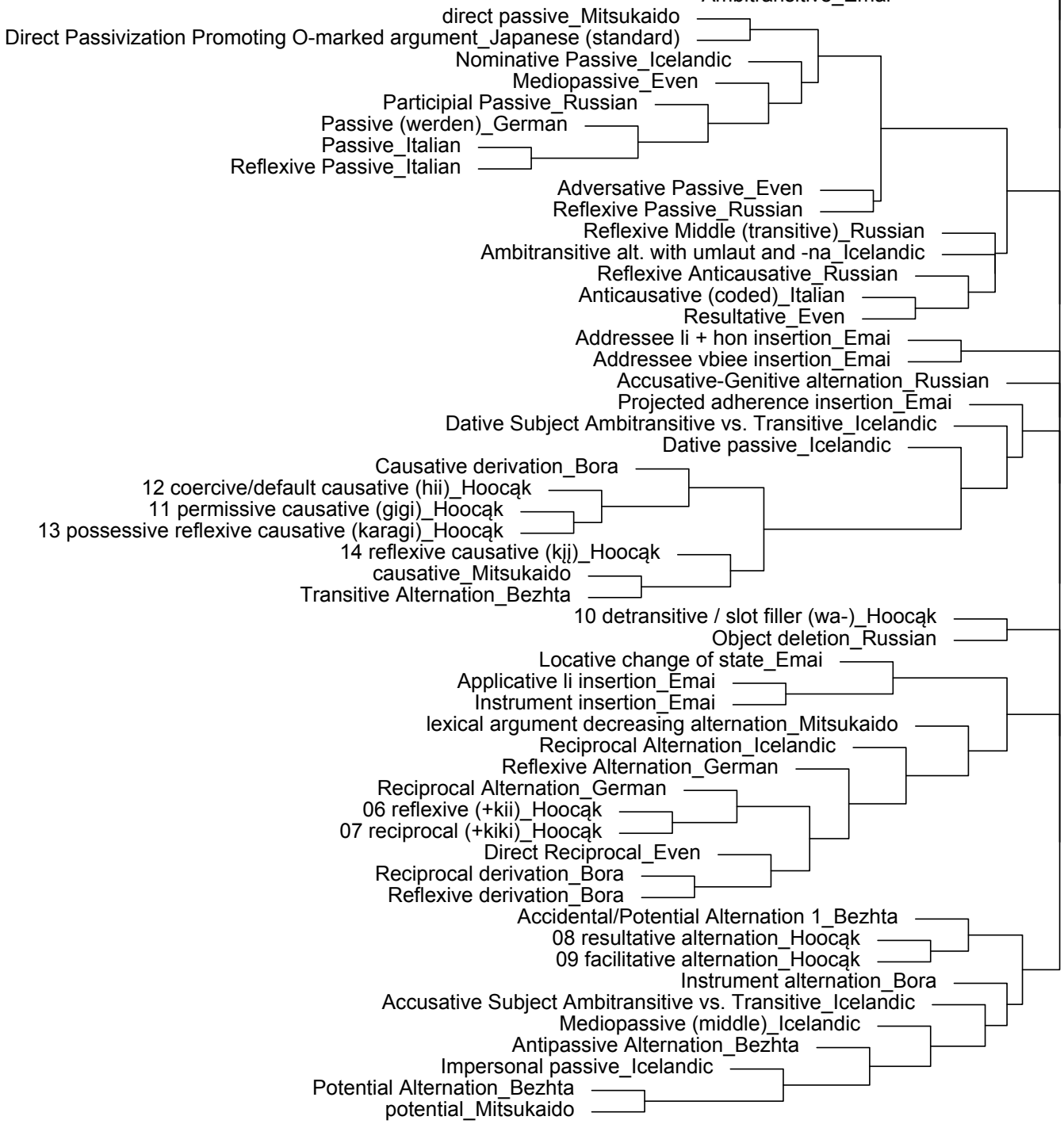
02 benefactive/possession of U (+gi)_Hoocąk



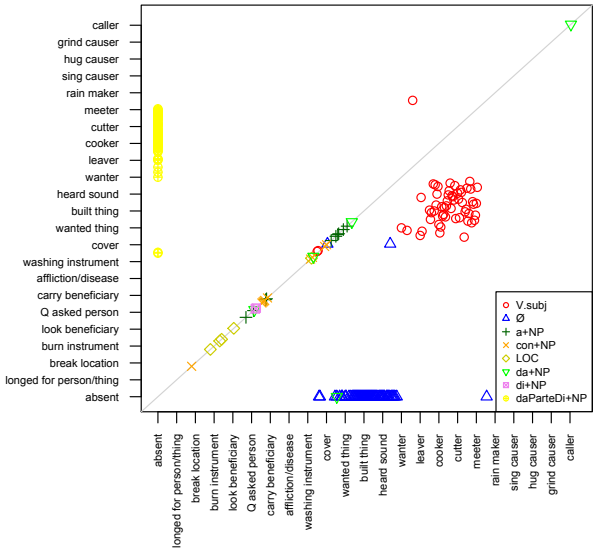
How many points are shared?

That is a combination of how the roles are linked between unmarked and marked alternant, but also of the lexical distribution of an alternation

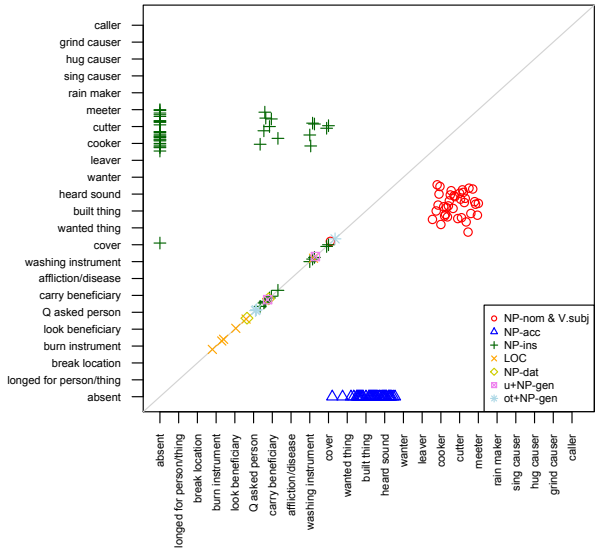

Transitive from intransitive_Icelandic
Reflexive Middle (intransitive)_Russian
Reciprocal Reflexive_Russian
ver- Alternation (I)_German
Prepositional Alternation (I)_German
Recipient Alternation_Bezhta
Locative Alternation_Japanese (standard)
lexical argument increasing alternation_Mitsukaido
Instrument Alternation_Bezhta
Locative omission_Emai
Impersonal mediopassive_Icelandic
Genitive Passive_Icelandic
er- Alternation_German
indirect passive_Mitsukaido
Direct Passivization Promoting Ni-marked Argument_Japanese (standard)
Indirect Reciprocal_Even
Direct Obj. vs. Prepositional Obj_Icelandic
Dat-Acc Ambiditransitive_Icelandic
Comitative alternation_Russian
Reflexive (mediopassive form)_Icelandic
Comitative alternation_Even
Impersonal Passive_German
Cognate Object Insertion_German
object insertion_Mitsukaido
Reflexive deleting alternation_Even
Cognate object alternation_Russian
Object deletion_Even
Object omission_Emai
Object Omission Alternation_German
Benefactive Alternation_German
Negative Accusative-Genitive alternation_Russian
Object Omission_Italian
Oblique Subject_Italian
Semantic reflexive_Russian
Locative Alternation_Italian
Locative Alternation with T.acc +með I.dat_Icelandic
be- Alternation (II)_German
Prefixal Goal-Instrumental alternation_Russian
be- Alternation (I)_German
Prefixal Source-promoting alternation_Russian
Middle_Icelandic
Anticausative alternation_German
Instrumental subject alternation_Russian
Ambitransitive_Icelandic
Ambitransitive Alternation_German
Ambitransitive Alternation_Bezhta
Caused-motion from intransitive_Icelandic
Elative insertion_Emai
Allative insertion_Emai
Ambitransitive_Emai
direct passive_Mitsukaido
Direct Passivization Promoting O-marked argument_Japanese (standard)
Nominative Passive_Icelandic
Mediopassive_Even
Participial Passive_Russian
Passive (werden)_German
Passive_Italian
Reflexive Passive_Italian
Adversative Passive_Even
Reflexive Passive_Russian
Reflexive Middle (transitive)_Russian
Ambitransitive alt. with umlaut and -na_Icelandic
Reflexive Anticausative_Russian
Anticausative (coded)_Italian
Resultative_Even
Addressee li + hon insertion_Emai
Addressee vbiee insertion_Emai
Accusative-Genitive alternation_Russian
Projected adherence insertion_Emai
Dative Subject Ambitransitive vs. Transitive_Icelandic
Dative passive_Icelandic
Causative derivation_Bora
12 coercive/default causative (hii)_Hoocąk
11 permissive causative (gigi)_Hoocąk
13 possessive reflexive causative (karagi)_Hoocąk
14 reflexive causative (kjj)_Hoocąk
causative_Mitsukaido
Transitive Alternation_Bezhta
10 detransitive / slot filler (wa-)_Hoocąk
Object deletion_Russian
Locative change of state_Emai
Applicative li insertion_Emai
Instrument insertion_Emai
lexical argument decreasing alternation_Mitsukaido
Reciprocal Alternation_Icelandic
Reflexive Alternation_German
Reciprocal Alternation_German
06 reflexive (+kii)_Hoocąk
07 reciprocal (+kiki)_Hoocąk
Direct Reciprocal_Even
Reciprocal derivation_Bora
Reflexive derivation_Bora
Accidental/Potential Alternation 1_Bezhta
08 resultative alternation_Hoocąk
09 facilitative alternation_Hoocąk
Instrument alternation_Bora
Accusative Subject Ambitransitive vs. Transitive_Icelandic
Mediopassive (middle)_Icelandic
Antipassive Alternation_Bezhta
Impersonal passive_Icelandic
Potential Alternation_Bezhta
potential_Mitsukaido


direct passive_Mitsukaido
Direct Passivization Promoting O-marked argument_Japanese (standard)
Nominative Passive_Icelandic
Mediopassive_Even
Participial Passive_Russian
Passive (werden)_German
Passive_Italian
Reflexive Passive_Italian
Adversative Passive_Even
Reflexive Passive_Russian
Reflexive Middle (transitive)_Russian
Ambitransitive alt. with umlaut and -na_Icelandic
Reflexive Anticausative_Russian
Anticausative (coded)_Italian
Resultative_Even
Addressee li + hon insertion_Emai
Addressee vbiee insertion_Emai
Accusative-Genitive alternation_Russian
Projected adherence insertion_Emai
Dative Subject Ambitransitive vs. Transitive_Icelandic
Dative passive_Icelandic
Causative derivation_Bora
12 coercive/default causative (hii)_Hoocąk
11 permissive causative (gigi)_Hoocąk
13 possessive reflexive causative (karagi)_Hoocąk
14 reflexive causative (kįį)_Hoocąk
causative_Mitsukaido
Transitive Alternation_Bezhta
10 detransitive / slot filler (wa-)_Hoocąk
Object deletion_Russian
Locative change of state_Emai
Applicative li insertion_Emai
Instrument insertion_Emai
lexical argument decreasing alternation_Mitsukaido
Reciprocal Alternation_Icelandic
Reflexive Alternation_German
Reciprocal Alternation_German
06 reflexive (+kii)_Hoocąk
07 reciprocal (+kiki)_Hoocąk
Direct Reciprocal_Even
Reciprocal derivation_Bora
Reflexive derivation_Bora
Accidental/Potential Alternation 1_Bezhta
08 resultative alternation_Hoocąk
09 facilitative alternation_Hoocąk
Instrument alternation_Bora
Accusative Subject Ambitransitive vs. Transitive_Icelandic
Mediopassive (middle)_Icelandic
Antipassive Alternation_Bezhta
Impersonal passive_Icelandic
Potential Alternation_Bezhta
potential_Mitsukaido

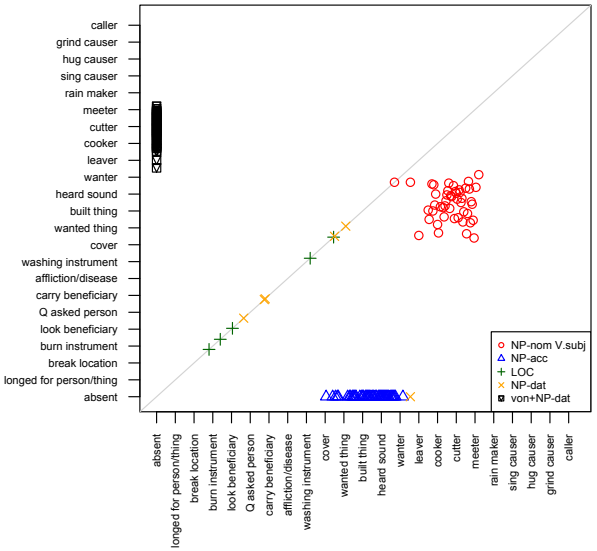
Passive from Italian

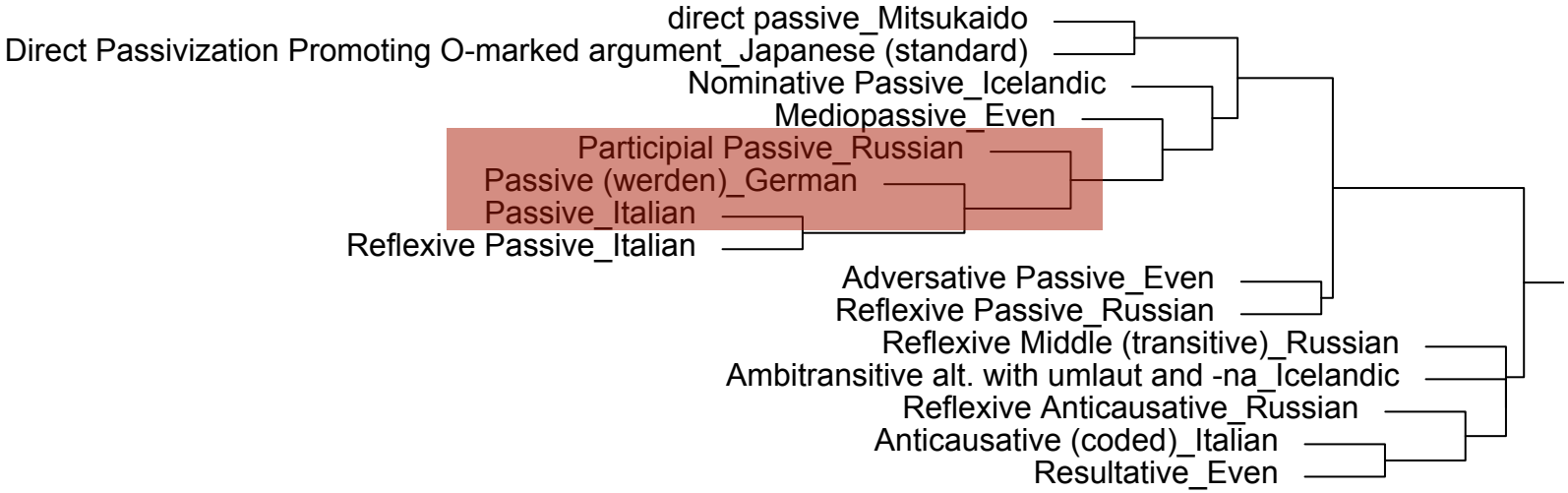


Participial Passive from Russian

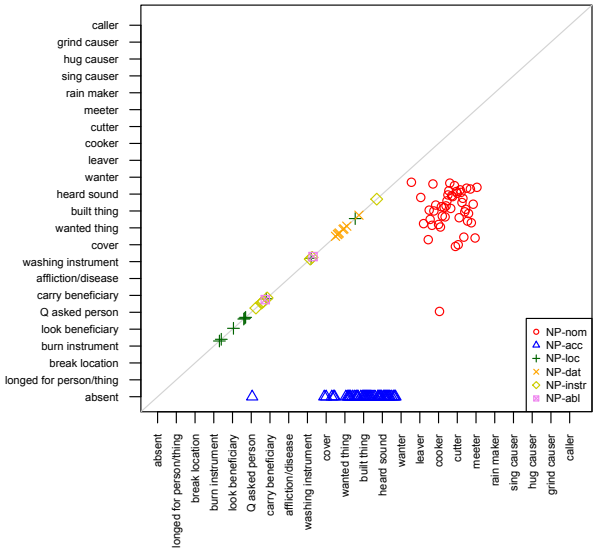


Passive (werden) from German

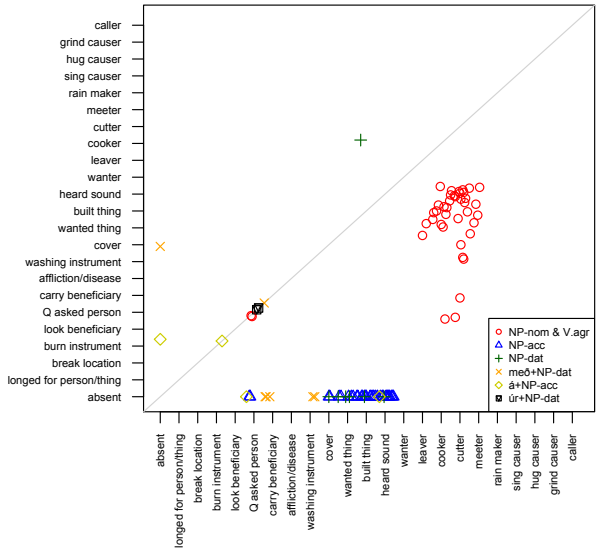




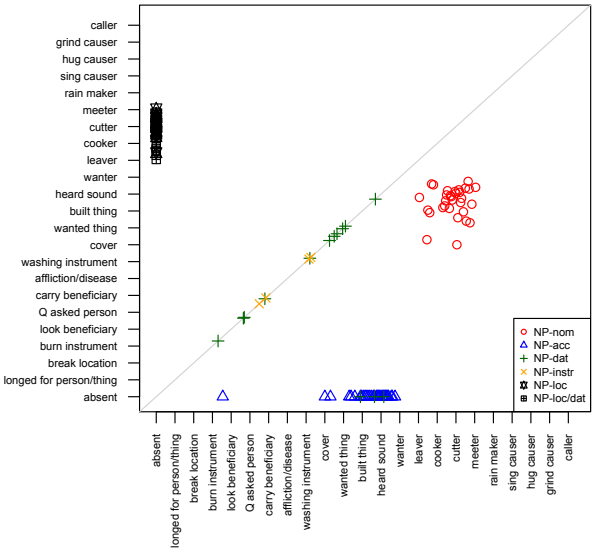
ssivization Promoting O-marked argument from Japane

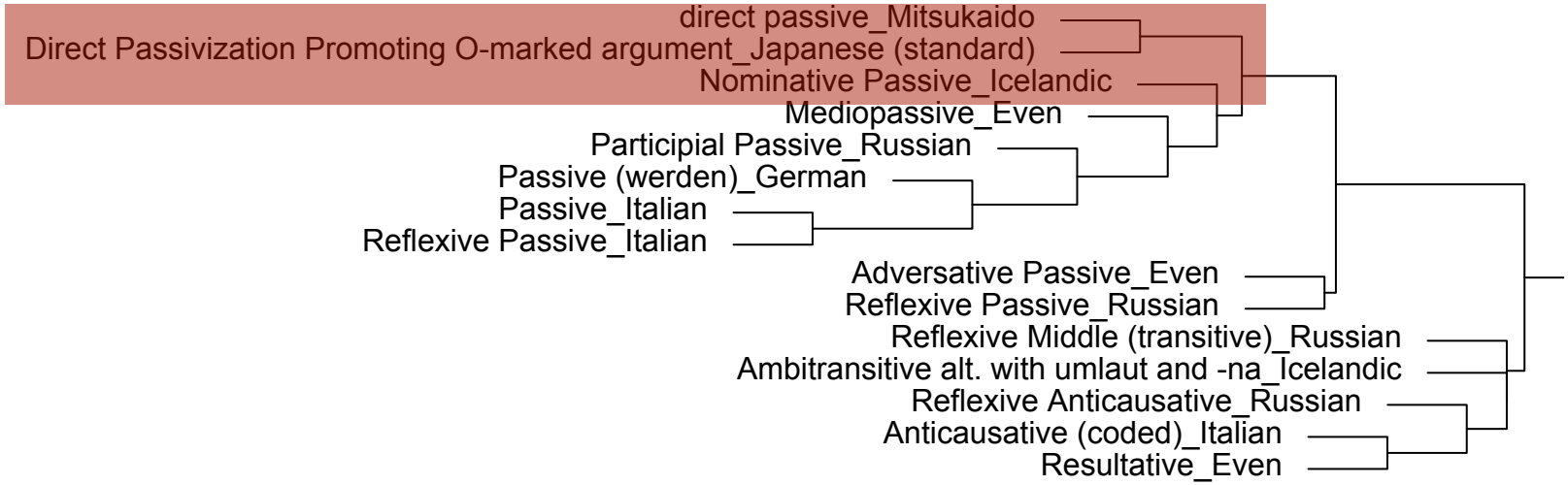


Nominative Passive_Icelandic

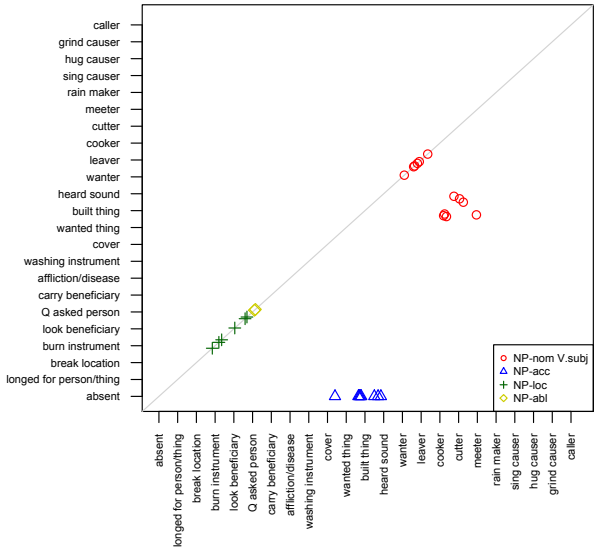


direct passive from Mitsukaido

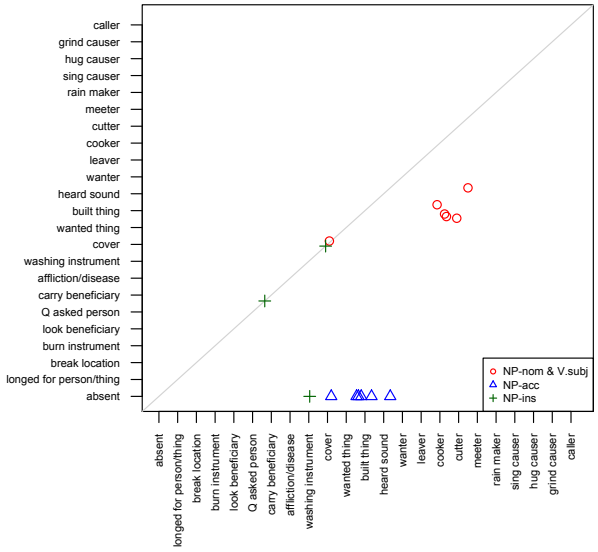




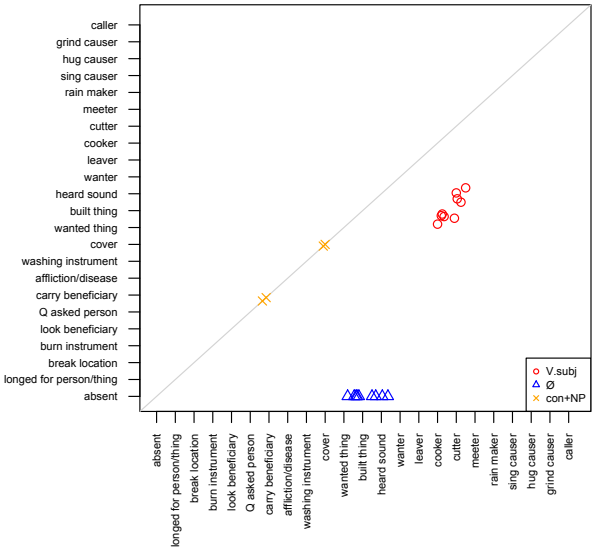
Resultative from Even

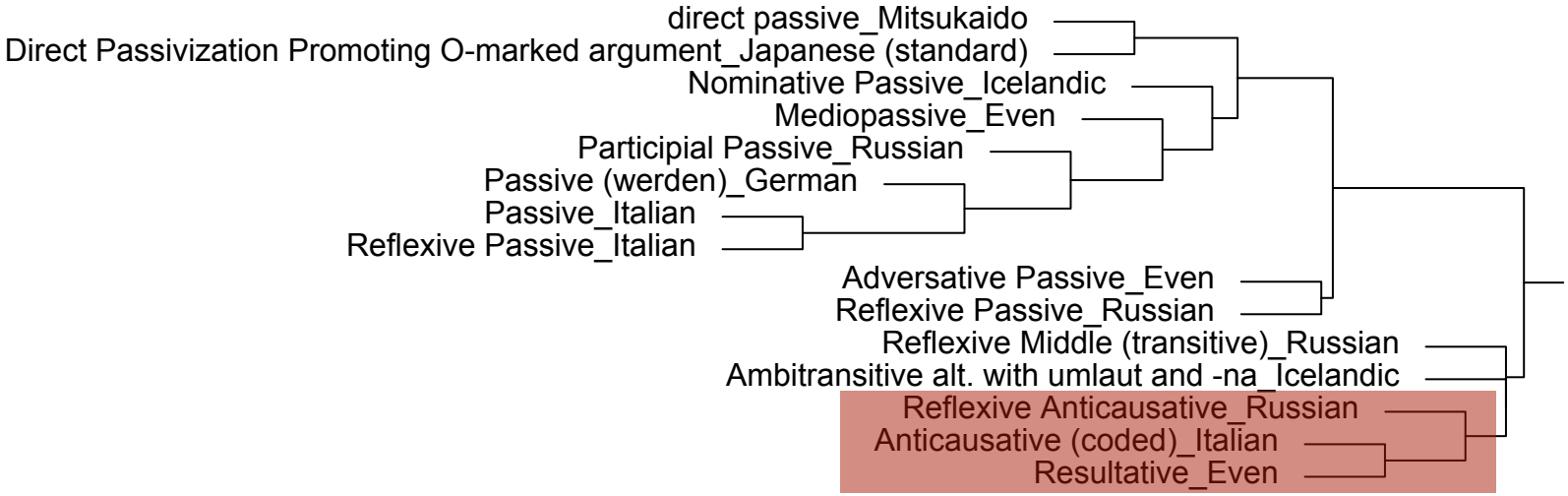


Reflexive Anticausative from Russian

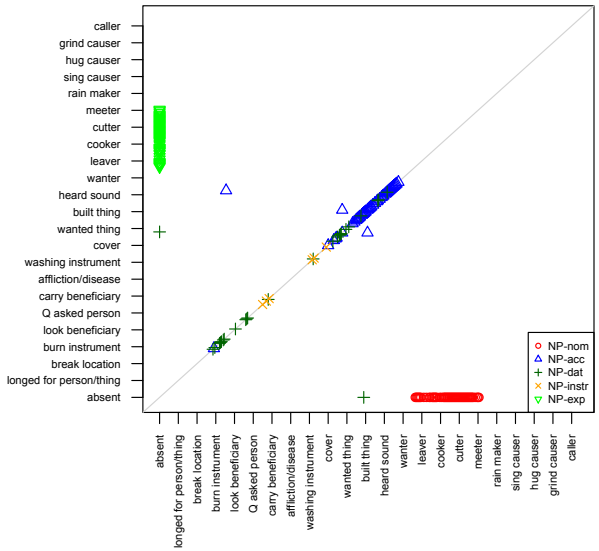


Anticausative (coded) from Italian

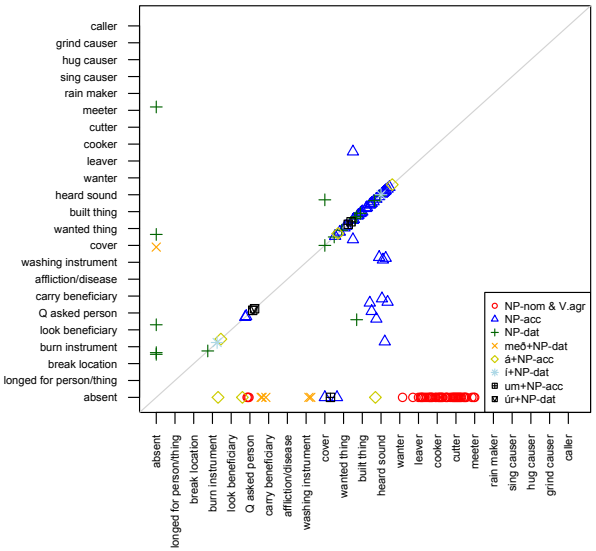




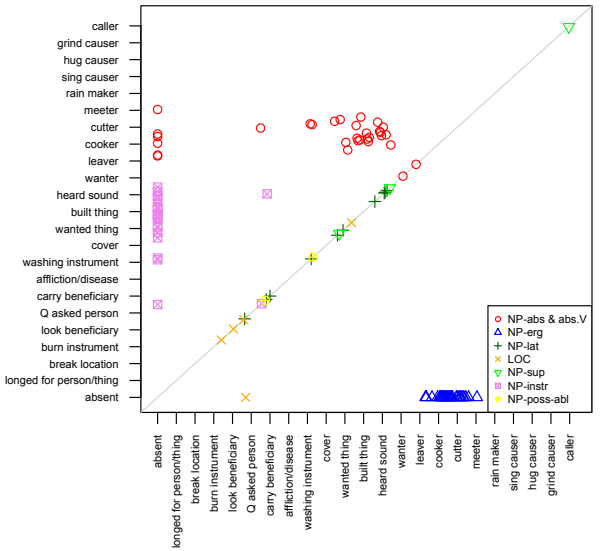
potential from Mitsukaido



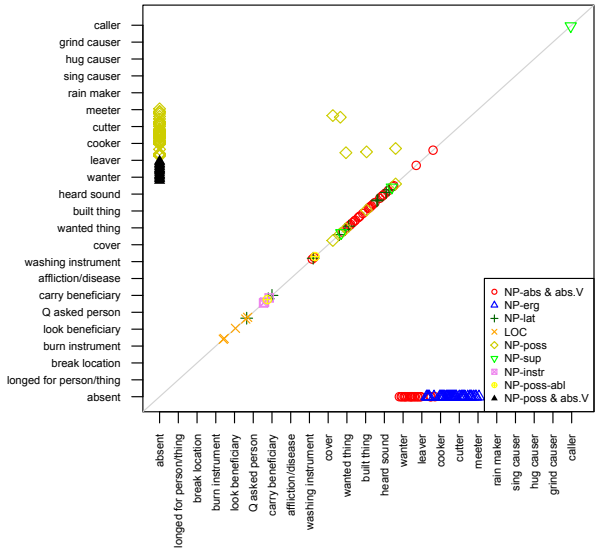
Impersonal passive_Icelandic

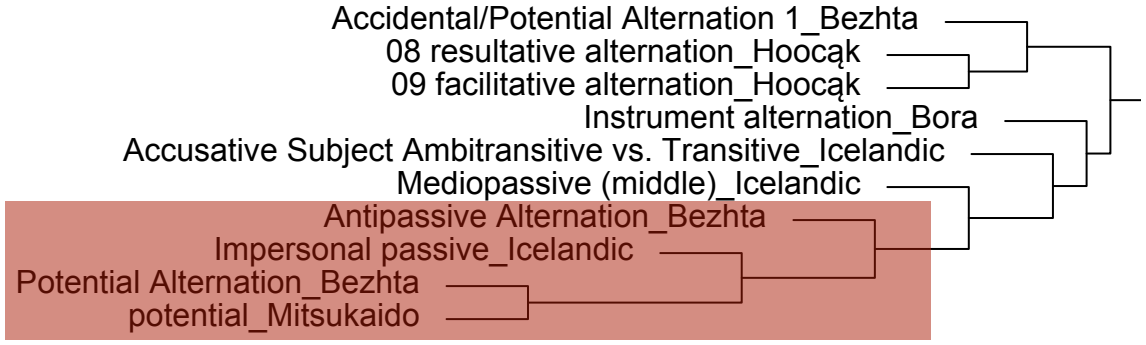


Antipassive Alternation from Bezhta

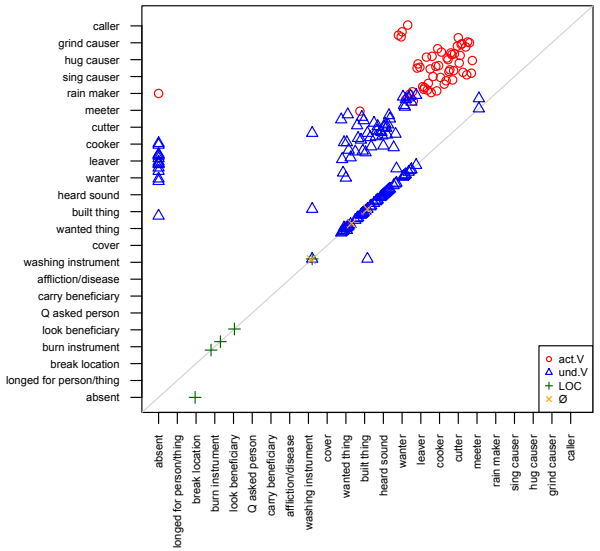


Potential Alternation from Bezhta

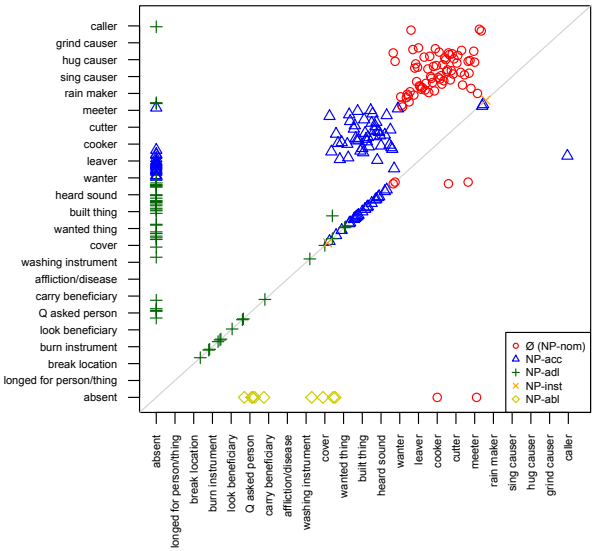


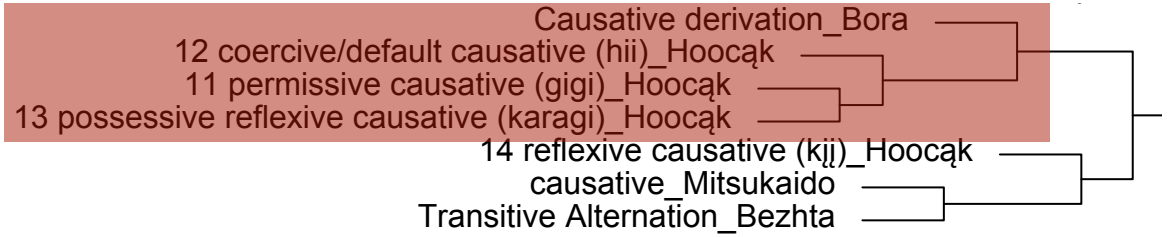


12 coercive/default causative (hii)_Hoocąk

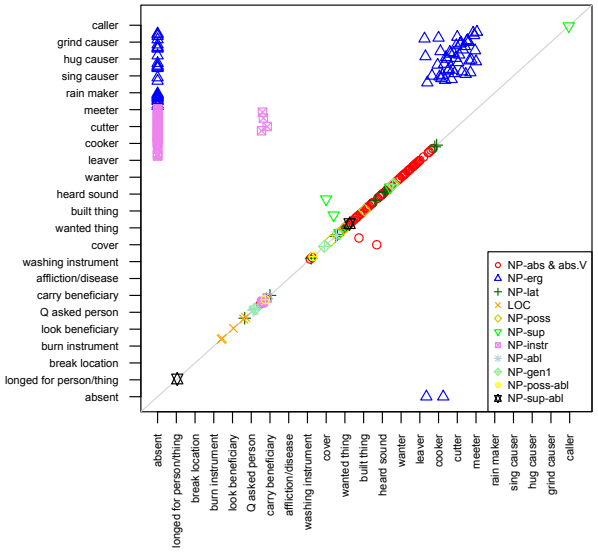


Causative derivation from Bora

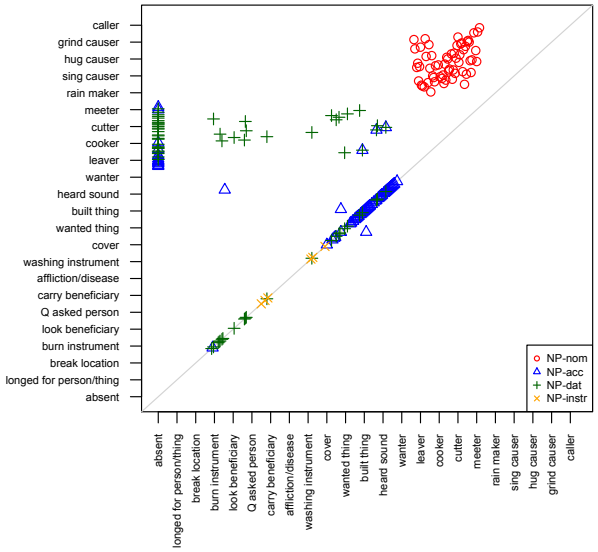




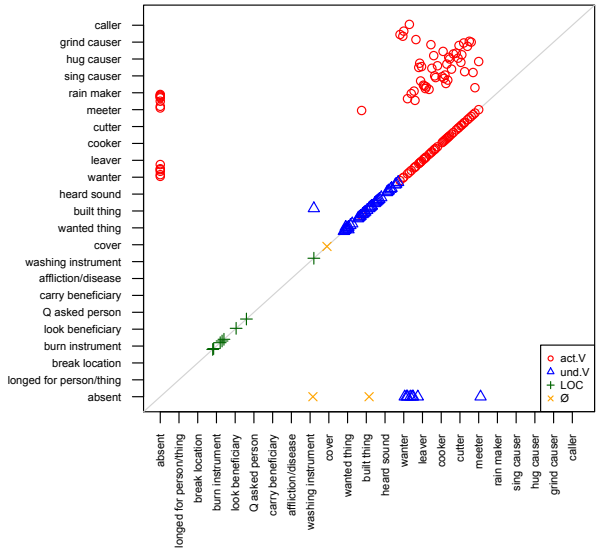
Transitive Alternation from Bezhta

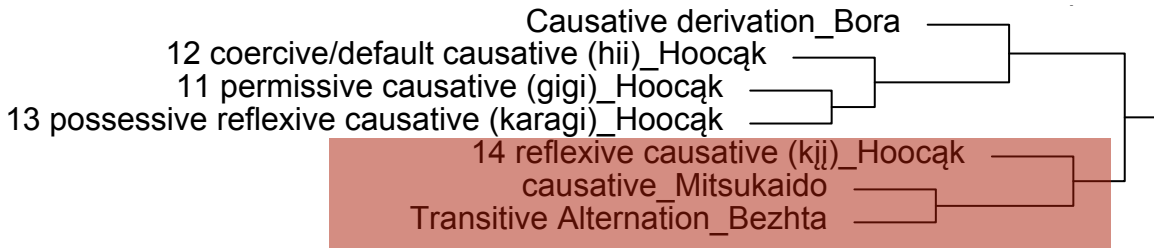


causative from Mitsukaido

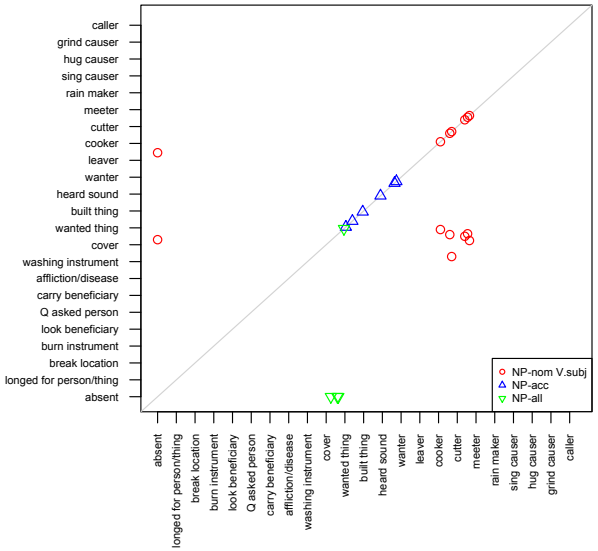


14 reflexive causative (kjj)_Hoocąk

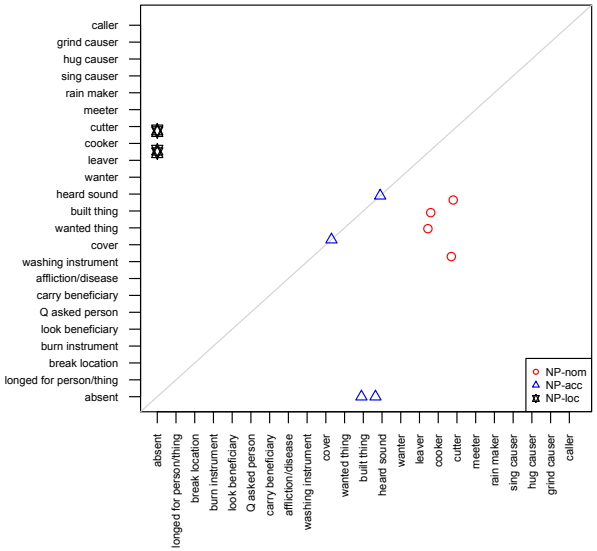




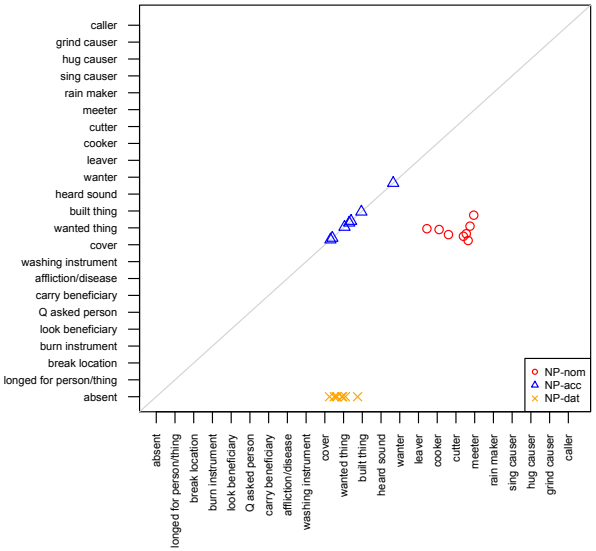
Indirect Reciprocal from Even



indirect passive from Mitsukaido



ssivization Promoting Ni-marked Argument from Japan



Summary

- Objective way of finding **alternation types**
- No typological comparison of language-specific coding necessary
- We would get exactly the same results if **all language-specific constructions were characterized by a random code**
- Likewise, **the actual meaning of the microroles is unimportant for this method**
- Based completely on **mapping between microroles and language-specific coding**