TYPOLOGICAL SURVEY OF TRANSITIVITY

Katya Aplonova
aplooon@gmail.com

OUTLINE

- Research objectives and hypothesis to be tested
- Previous researches
- Input data, descriptive statistics, data visualization
- Obtained results and their linguistic interpretation

RESEARCH OBJECTIVES AND HYPOTHESIS TO BE TESTED

TRANSITIVITY

- transitive construction is a basic construction with bivalent predicate

(1) *Петя бросил камень.*

- set of values of semantic parameters that is universally associated with high transitivity:

  - actionality (as opposed to stativity)
  - telicity
  - volitionality
  - control of one participant of the other
  - …

[Hopper and Thompson's 1980]

INTRANSITIVITY

- intrasitive constructions are much more diverse

(2) *У Пети болит голова.*

(3)	*Pet'a-n'*	*m'el'-s*	*tukšn-i*	*t'e*	*panar-os'*
	Petja-GEN	mind-ILL	go-PRS.3SG	this	shirt-NOM.DEF

'Petja likes this shirt'.

(4)	*Bhǎnkùlǎ*	*ĺĭǎ*	*à*	*wàà*	*Tíá*	*gí.*
	Макура	любовь	3SG.PRF	прибывать	Тиа	в

'Тиа влюбился в Макуру'.

ORIGINAL DATA

S. Say's database, very complex

	Предикат		чувства	русский: предикат	русский язык: X	русский язык: Y	Русский: Локус	РЯ: комментарий	русский: упрощенно	арабский: глагол	арабский: X	арабский: Y	арабский: локус	арабский: комментарий	арабский: упрощенные: кодирование	гуарани: глагол	гуарани X	гуарани Y	гуарани: локус	гуарани: комментарии	гуарани: упрощенный
1	болеть-1	у X-а (болеть) Y			uGEN	NOM	X		uGENNOM					NV	*						
2	болеть-2	X (болеть) Y-ом			NOM	INS	Y		NOMINS		NOM	bi	Y	NV	N_bi						
3	бояться	X (бояться) Y-а	]		NOM	GEN	Y		NOMGEN		NOM	min	Y		N_min		act	gui			actgui
4	бросить	X (бросить) Y(-а)			TR		TR		TR		TR		TR		TR		TR				TR
5	быть достаточно	X-у (быть достаточно) Y-а		хватать	DAT	GEN	XY		DATGEN		li	NOM	X	NV	li_N		inact	pe			inactpe
6	быть похожим	X (быть похожим) на У(-а)			NOM	naACC	Y	NV	NOMnaACC		TR		TR		TR					также 'быть равным кому/чему'; прим	
7	верить	X (верить)Y-у			NOM	DAT	Y		NOMDAT		TR		TR		TR		act	rehe			actrehe
8	взять	X (брать) Y(-а)			TR		TR		TR		TR		TR		TR		TR			примеры только с неодуш. 2-м	TR
9	видеть	X (видеть) Y(-а)			TR		TR		TR		TR		TR		TR		TR				TR
10	влиять	X (влиять) на Y(-а)			NOM	naACC	Y		NOMnaACC		NOM	alaa	Y		N_'alaa						
11	встречаться	X (встречаться) с Y-ом			NOM	sINS	Y		NOMsINS		NOM	bi	Y		N_bi		TR			1) = 55; также 'встречать кого,	TR
12	входить	X (входить) в Y (дом)			NOM	vACC	Y		NOMvACC		TR		TR		TR		act	pe		1) рамка управления с Y-re – в	actpe(lo
13	выиграть	X (выигрывать) у Y-а			NOM	uGEN	Y		NOMuGEN						*						
14	выходить	X (выходить) из Y-а			NOM	izGEN	Y		NOMizGEN		NOM	min	Y		N_min		act	gui			actgui
15	гнать	X (гнать) Y-а			TR		TR		TR		TR		TR		TR		TR			'убирать, вынимать', может та	TR
16	гнуть	X (гнуть) Y			TR		TR		TR		TR		TR		TR					нет примеров	
17	говорить	X (говорить) Y-у			NOM	DAT	Y		NOMDAT		NOM	li	Y		N_li					имеет нерегулярное спряжение	
18	держать	X (держать) Y(-а)			TR		TR		TR		TR		TR		TR		TR			также 'хватать, останавливать'	TR
19	догнать	X (догонять) Y(-а)			TR		TR		TR		NOM	bi	Y		N_bi		TR			21	TR

LANGUAGES

Russian, Standart Arabic, Guarani, Estonian, Tsaxur, Tuvan, Ingermanlandic, Basque, French, German, Bagwalal, Japanese, Lithuanian, Kalmyk, Khmer, Bashkir, Latvian, Guro, Looma, Lezgi, Modern Greek, Ancient Greek, Albanian, Spanish, Irish, Armenian, Azerbaijani, Romani (Kalderash), English, Mandarin, Polish, Dutch, Italian, Komi-Zyrian, Ossetic, Serbian, Chukchee, Norwegian (B).

TOTAL: 38 languages

LANGUAGES

Russian, Standart Arabic, Guarani, Estonian, Tsaxur, Tuvan, Ingermanlandic, Basque, **French**, **German**, Bagwalal, Japanese, **Lithuanian**, Kalmyk, Khmer, Bashkir, **Latvian**, Guro, Looma, Lezgi, **Modern Greek**, **Ancient Greek**, **Albanian**, **Spanish**, **Irish**, **Armenian**, Azerbaijani, **Romani (Kalderash)**, **English**, Mandarin, **Polish**, **Dutch**, **Italian**, Komi-Zyrian, **Ossetic**, **Serbian**, Chukchee, **Norwegian (B)**.

Indo-European: 19 languages

CONSTRUCTIONS

feel_pain, have#illness#, be_afraid, throw, have_enough, resemble, believe, take, see, influence, encounter, enter, win, go_out, drive, bend, tell, hold, catch_up, milk, reach, touch, fight, be_friends, think, eat, fry, wait, forget,…

TOTAL: 130 constructions

RESEARCH QUESTIONS

- Do languages/constructions form distinctive groups depending on (in)transitivity?
- What are the outliers? Why?
- Do areal and genetic characteristics metter?

PREVIOUS RESEARCHES

were made by S. Say
(using the same database)

BIVALENT VERB CLASSES IN THE LANGUAGES OF EUROPE: A QUANTITATIVE TYPOLOGICAL STUDY

Objectives:

- to propose methods for measuring (dis)similarities in the organization of valency class systems across languages
- to test them on a sample of European languages in order to reveal areal and genetic patterns

Main conclusion:

- (in)transitivity is correlated with areal factors

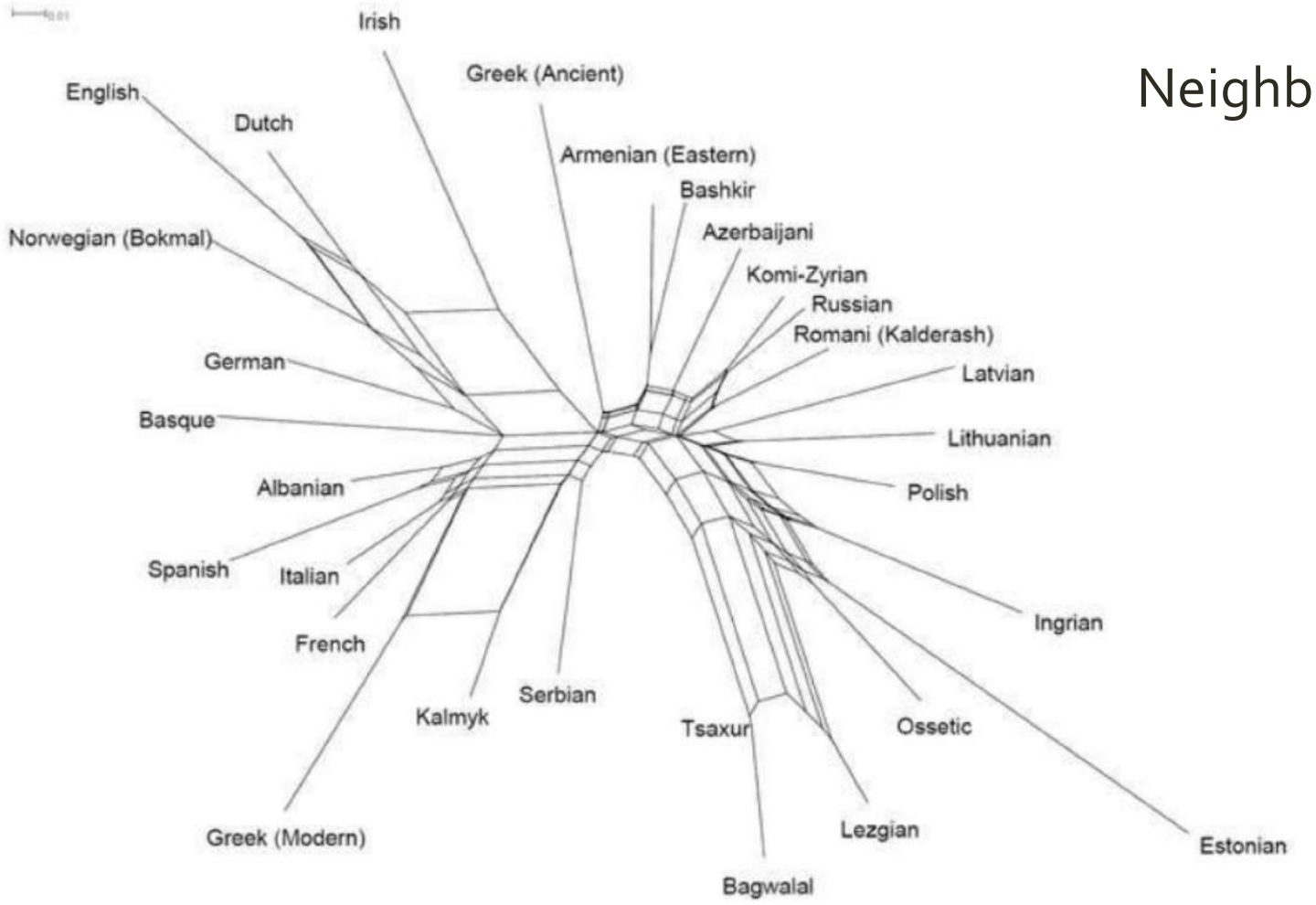
[Say 2014]

METHODS OF ANALYSIS

- comparison of (in)transitivity ratio
- Pearson's correlation for intrasitivity ration and number of cases
- degree of dissimilarity in pairs of languages (relative Hamming distance) : the proportion of predicates with non-identical transitivity values out of all predicates that have been obtained for both languages

[Say 2014]

EXAMPLE OF VISUALIZATION



NeighbourNet

INPUT DATA,DATA VISUALIZATION

BASIC DATAFRAMES

- **df**: categorical, languages as variables

	Construction <fctr>	Russian <fctr>	Arabic <fctr>	Guarani <fctr>	Estonian <fctr>	Tsaxur <fctr>	Tuvan <fctr>	Ingerm <fctr>	Basque <fctr>	French <fctr>	German <fctr>	Bagwalal <fctr>	Japanese <fctr>	Lithuanian <fctr>	Kalmyk <fctr>
1	feel_pain	intr	unknown	unknown	intr	unknown	intr	intr	unknown	intr	intr	intr	intr	intr	unknown
2	have#illness#	intr	intr	unknown	unknown	unknown	unknown	intr	tr	tr	intr	intr	intr	intr	intr
3	be_afraid	intr	intr	intr	intr	intr	intr	intr	intr	intr	intr	intr	unknown	intr	intr
4	throw	tr	tr	tr	unknown	tr	tr	tr	tr	tr	tr	tr	tr	tr	tr
5	have_enough	intr	intr	intr	intr	intr	intr	intr	unknown	unknown	intr	unknown	intr	intr	intr
6	resemble	intr	tr	unknown	intr	unknown	intr	unknown	unknown	intr	intr	intr	intr	intr	tr

6 rows | 1–16 of 43 columns

- **tdf**: categorical, constructions as variables

	Language <fctr>	have.illness. <fctr>	be_afraid <fctr>	throw <fctr>	have_enough <fctr>	resemble <fctr>	believe <fctr>	take <fctr>	see <fctr>	influence <fctr>	encounter <fctr>	enter <fctr>	win <fctr>	go_outr <fctr>	drive <fctr>
1	Russian	intr	intr	tr	intr	intr	intr	tr	tr	intr	intr	intr	intr	intr	tr
2	Arabic	intr	intr	tr	intr	tr	tr	tr	tr	intr	intr	tr	unknown	intr	tr
3	Guarani	unknown	intr	tr	intr	unknown	intr	tr	tr	unknown	tr	intr	unknown	intr	tr
4	Estonian	unknown	intr	unknown	intr	intr	unknown	tr	intr	intr	intr	intr	tr	intr	tr
5	Tsaxur	unknown	intr	tr	intr	unknown	unknown	tr	intr	unknown	intr	intr	intr	intr	tr
6	Tuvan	unknown	intr	tr	intr	intr	intr	tr	tr	intr	intr	intr	tr	intr	tr

6 rows | 1–16 of 132 columns

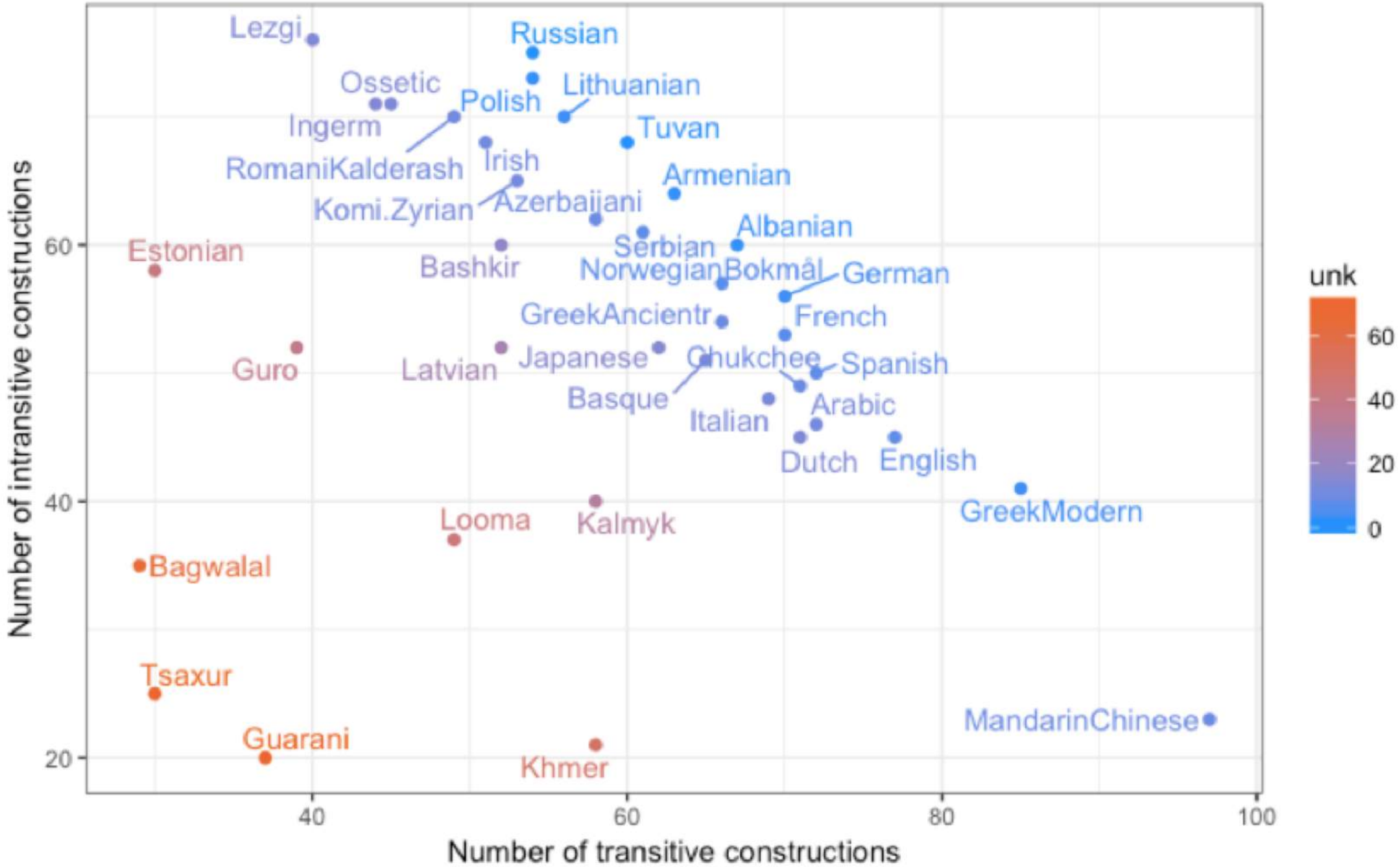
DATAFRAMES MADE FOR EXPLORATORY ANALYSIS

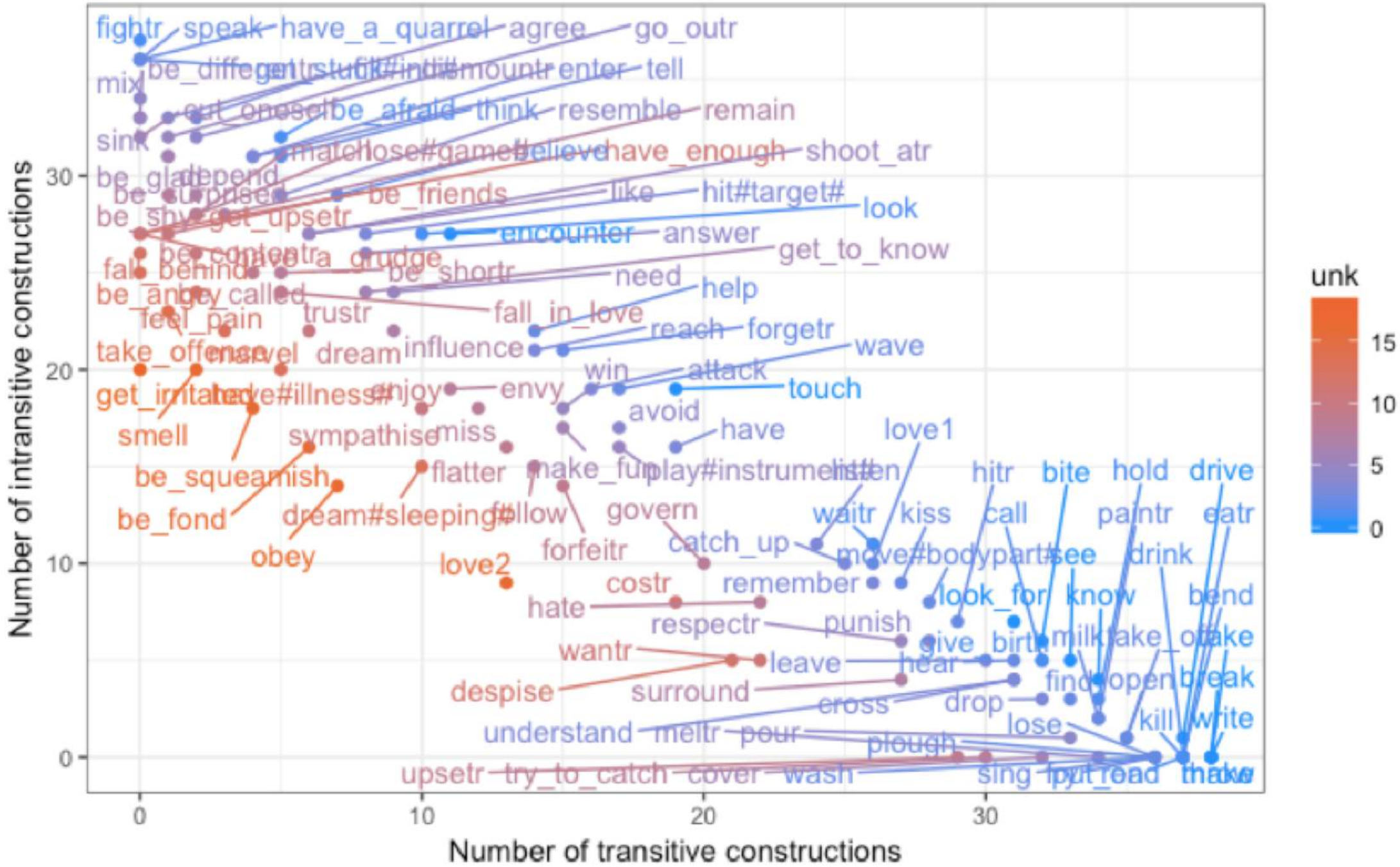
- **num**: numeric, languages as variables

```
        lang          tr             intr           unk
Albanian   : 1   Min.   :29.00   Min.   :20.00   Min.   : 0.00
Arabic     : 1   1st Qu.:49.50   1st Qu.:45.25   1st Qu.: 6.25
Armenian   : 1   Median :58.00   Median :53.50   Median :10.50
Azerbaijani: 1   Mean   :57.95   Mean   :53.26   Mean   :17.79
Bagwalal   : 1   3rd Qu.:68.50   3rd Qu.:64.75   3rd Qu.:16.50
Bashkir    : 1   Max.   :97.00   Max.   :76.00   Max.   :74.00
(Other)    :32
```

- **tnum**: numeric, constructios as variables

```
      constr          tr             intr           unk
agree    :  1   Min.   : 0.00   Min.   : 0.00   Min.   : 0.000
answer   :  1   1st Qu.: 4.00   1st Qu.: 4.25   1st Qu.: 2.000
attack   :  1   Median :14.50   Median :17.00   Median : 4.000
avoid    :  1   Mean   :16.95   Mean   :15.75   Mean   : 5.292
be_afraid:  1   3rd Qu.:31.00   3rd Qu.:27.00   3rd Qu.: 8.000
be_angry :  1   Max.   :38.00   Max.   :37.00   Max.   :18.000
(Other)  :124
```


Lezgi
Russian
Ossetic
Lithuanian
Polish
Ingerm
Tuvan
RomaniKalderash
Irish
Armenian
Komi.Zyrian
Azerbaijani
Albanian
Estonian
Serbian
Bashkir
Norwegian
Bokmål
German
GreekAncientr
French
Guro
Latvian
Japanese
Chukchee
Spanish
Basque
Arabic
Italian
Dutch
English
GreekModern
Looma
Kalmyk
Bagwalal
Tsaxur
MandarinChinese
Guarani
Khmer
Number of intransitive constructions
Number of transitive constructions
20
40
60
40
60
80
100
unk
0
20
40
60

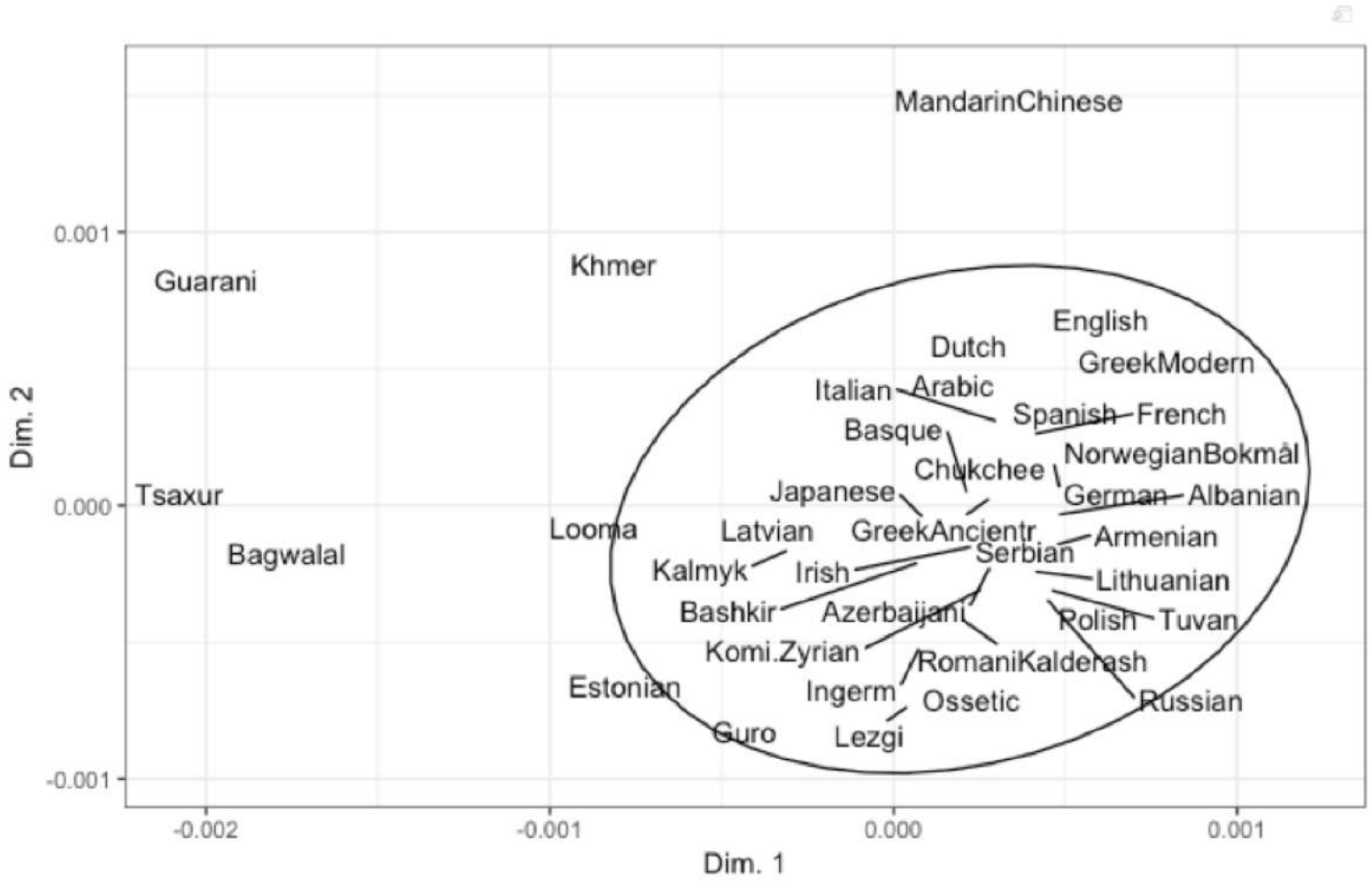

Number of intransitive constructions
Number of transitive constructions
0
10
20
30
unk
15
10
5
0
fightr
speak
have_a_quarrel
agree
go_outr
be_different
enter
tell
mix
think
resemble
remain
sink
be_afraid
have_enough
shoot_atr
be_glad
depend
be_surprised
be_friends
like
hit#target#
look
be_shy
get_upsetr
encounter
answer
get_to_know
be_contentr
bear_a_grudge
be_shortr
need
help
fall_behind
be_angry
be_called
feel_pain
trustr
fall_in_love
reach
forgetr
take_offence
travel
dream
influence
wave
win
attack
get_irritated
enjoy
envy
touch
smell
sympathise
miss
avoid
have
love1
be_squeamish
flatter
play#instrument#
hitr
bite
hold
drive
be_fond
dream#sleeping#
govern
waitr
kiss
call
paintr
eatr
obey
forfeitr
catch_up
love2
drink
costr
remember
look_for
know
bend
hate
respectr
punish
give
birth
milk
take_off
take
wantr
leave
hear
despise
surround
find
open
break
cross
drop
lose
understand
meltr
pour
kill
write
plough
upsetr
try_to_catch
cover
wash
sing
put
read
thank

SOME OBSERVATIONS

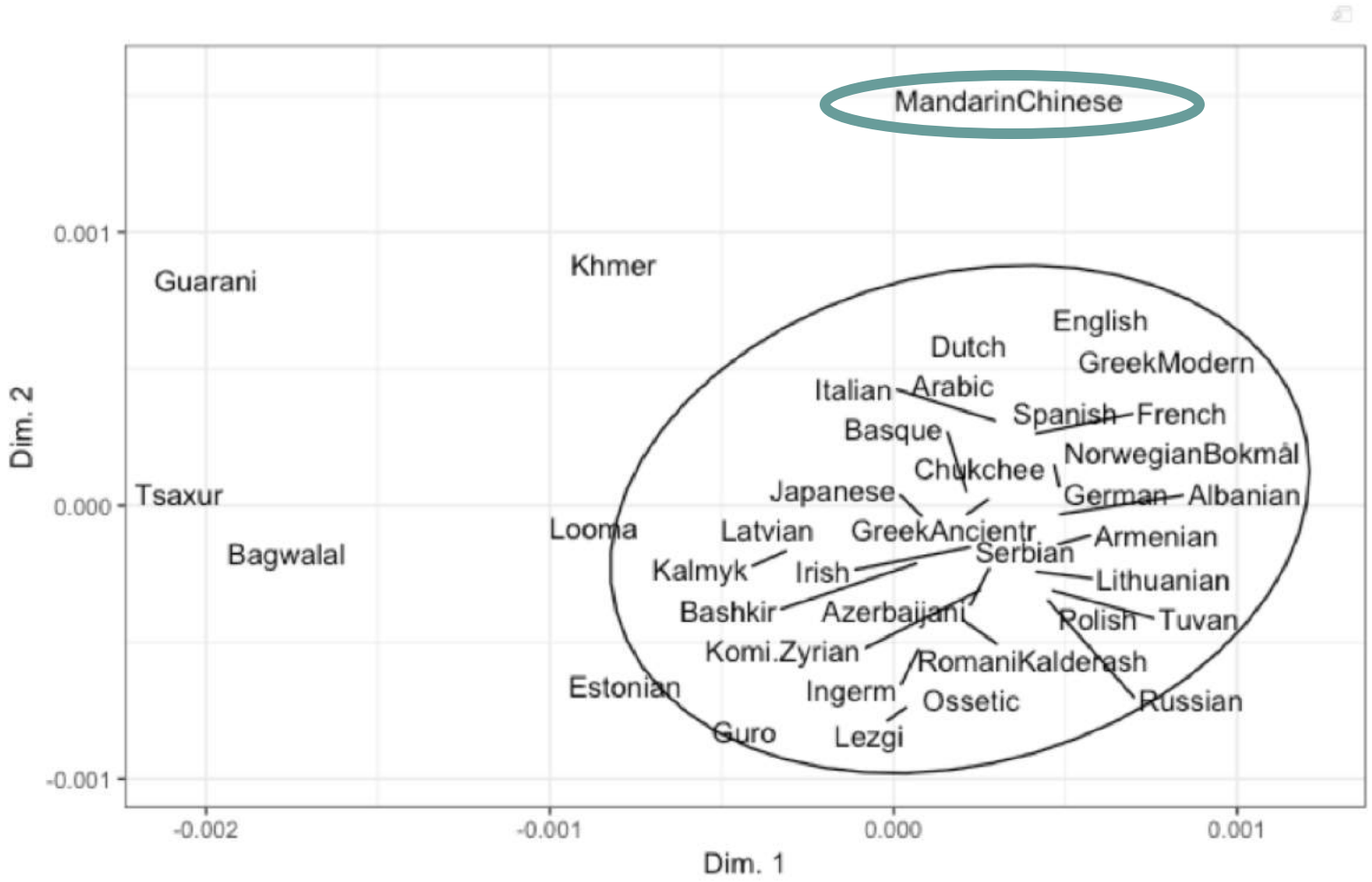
- There is a significant difference in distribution of (in)transitive constructions across languages. As far as constuctions are concerned, they do demonstrate prefernces to be (in)transitive.
- We need to solve a problem of empty values. On the plots the number of empty values for each variable is indicated by color. I see three possible solutions:
  - exclude from analysis languges and constructions which do not have enough data
  - normalize data
  - **make component analysis and exclude empty values from resultive plots (but is it really a solution?..)**

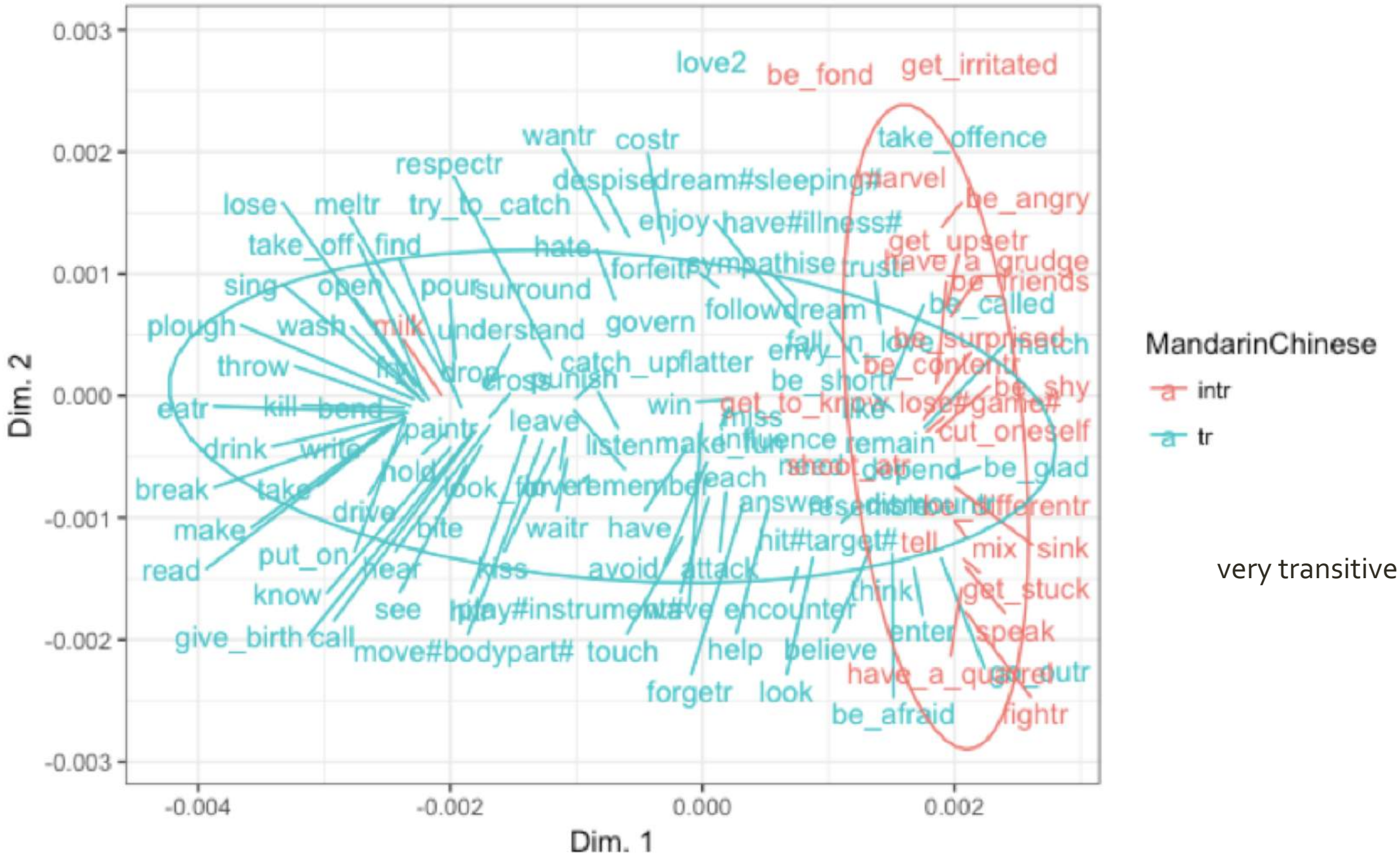
RESULTS AND THEIR LINGUISTIC INTERPRETATION

MCA FOR LANGUAGES

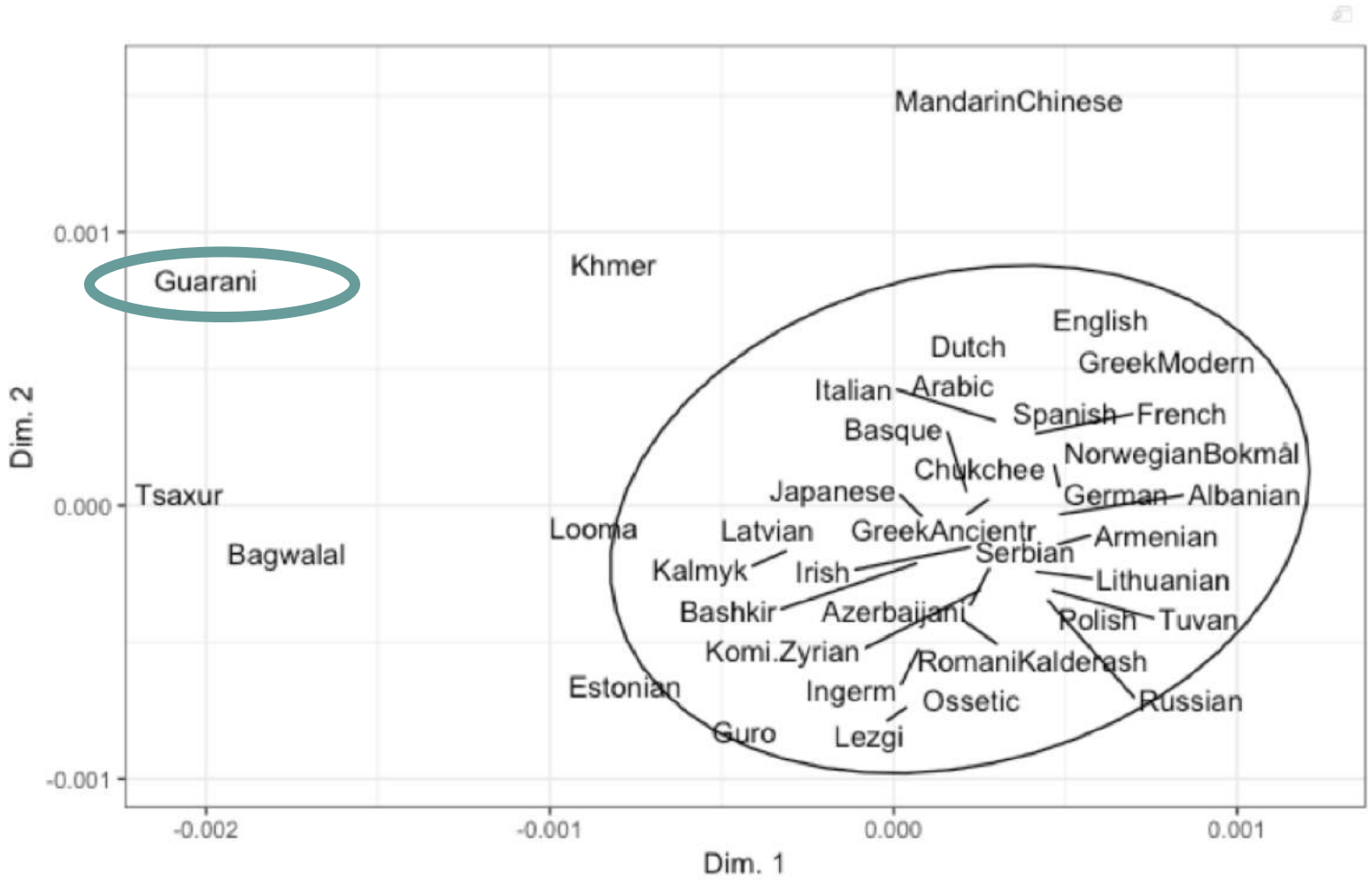


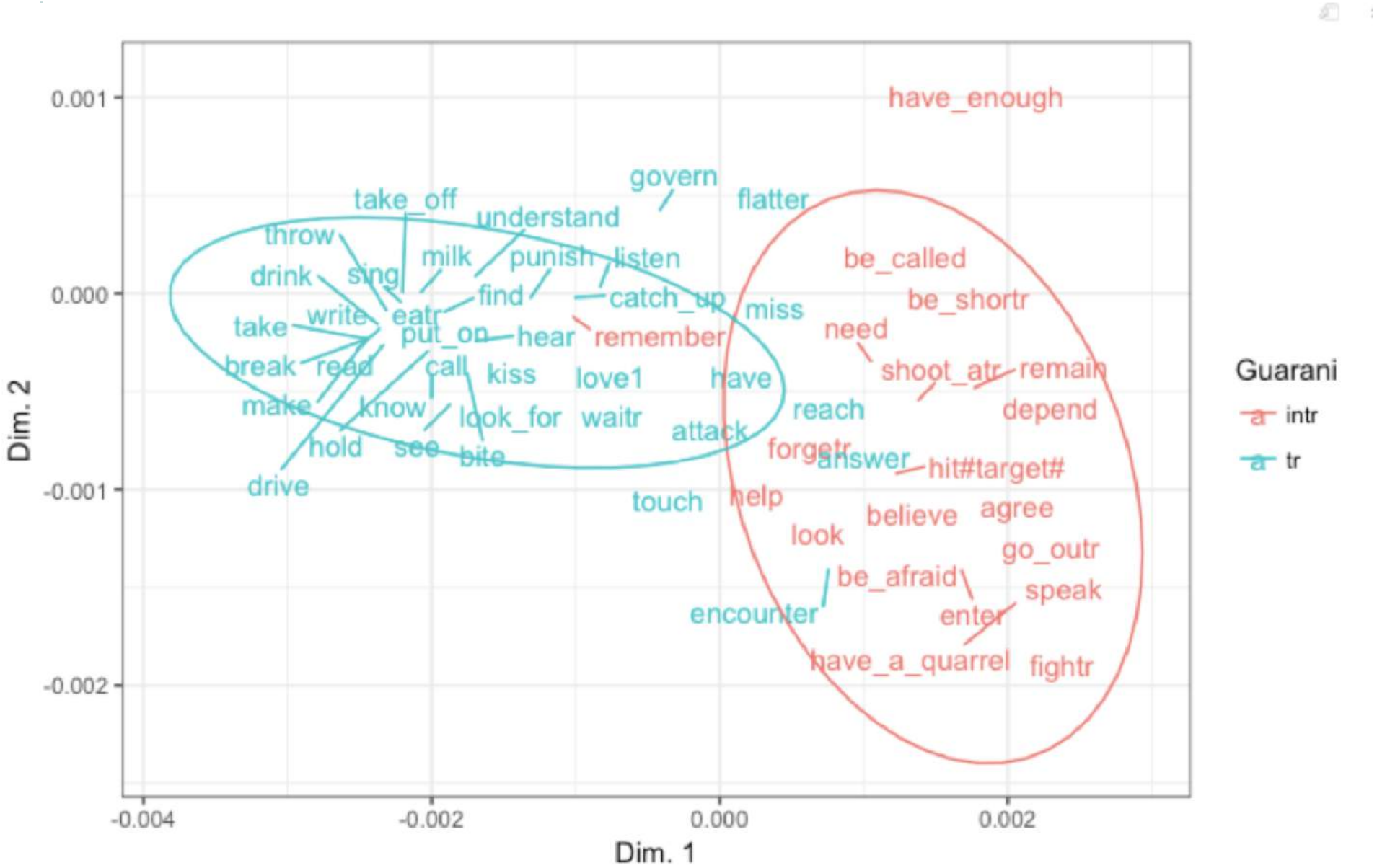
LET'S HAVE A LOOK AT OUTLIERS

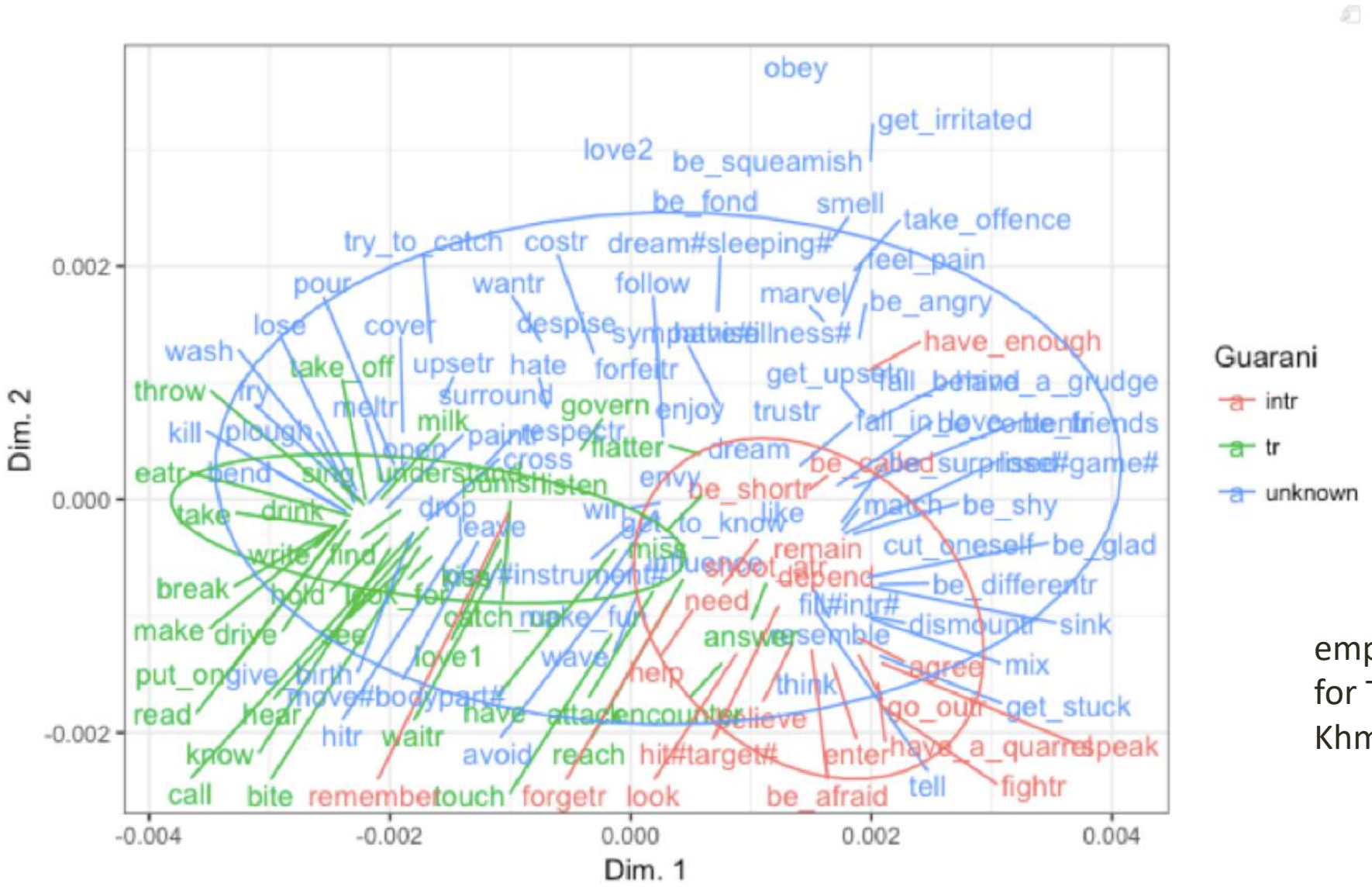



Dim. 2
Dim. 1
0.003
0.002
0.001
0.000
-0.001
-0.002
-0.003
-0.004
-0.002
0.000
0.002
love2
be_fond
get_irritated
take_offence
wantr
costr
respectr
despise
dream#sleeping#
marvel
be_angry
lose
meltr
try_to_catch
enjoy
have#illness#
get_upsetr
take_off
find
hate
forgetr
sympathise
trustr
have_a_grudge
be_friends
sing
open
pour
surround
be_called
plough
wash
milk
understand
govern
followdream
be_surprised
throw
catch_up
flatter
envy
be_contentr
eatr
kill
bend
drop
cross
punish
be_shortr
win
get_to_know
lose#game#
be_shy
paintr
leave
miss
like
drink
write
listen
make_fun
influence
remain
cut_oneself
hold
depend
be_glad
break
take
look_for
remember
teach
answer
resemble
different
drive
bite
waitr
have
hit#target#
tell
mix
sink
make
put_on
hear
kiss
avoid
attack
read
know
think
get_stuck
see
play#instrument#
have
encounter
enter
speak
give_birth
call
move#bodypart#
touch
help
believe
have_a_quarrel
worry
forgetr
look
be_afraid
fightr
ManderinChinese
intr
tr
very transitive

LET'S HAVE A LOOK AT OUTLIERS

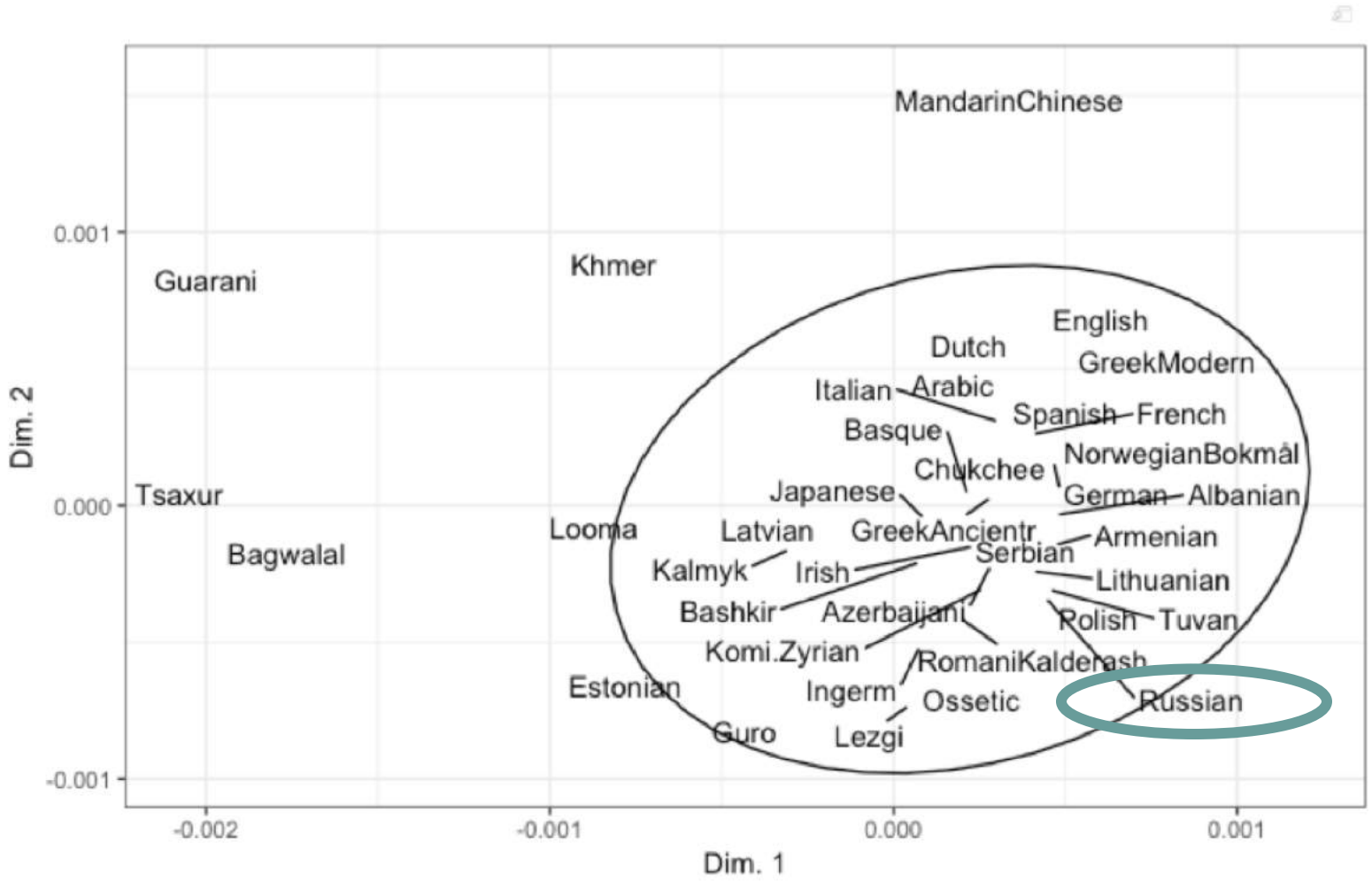


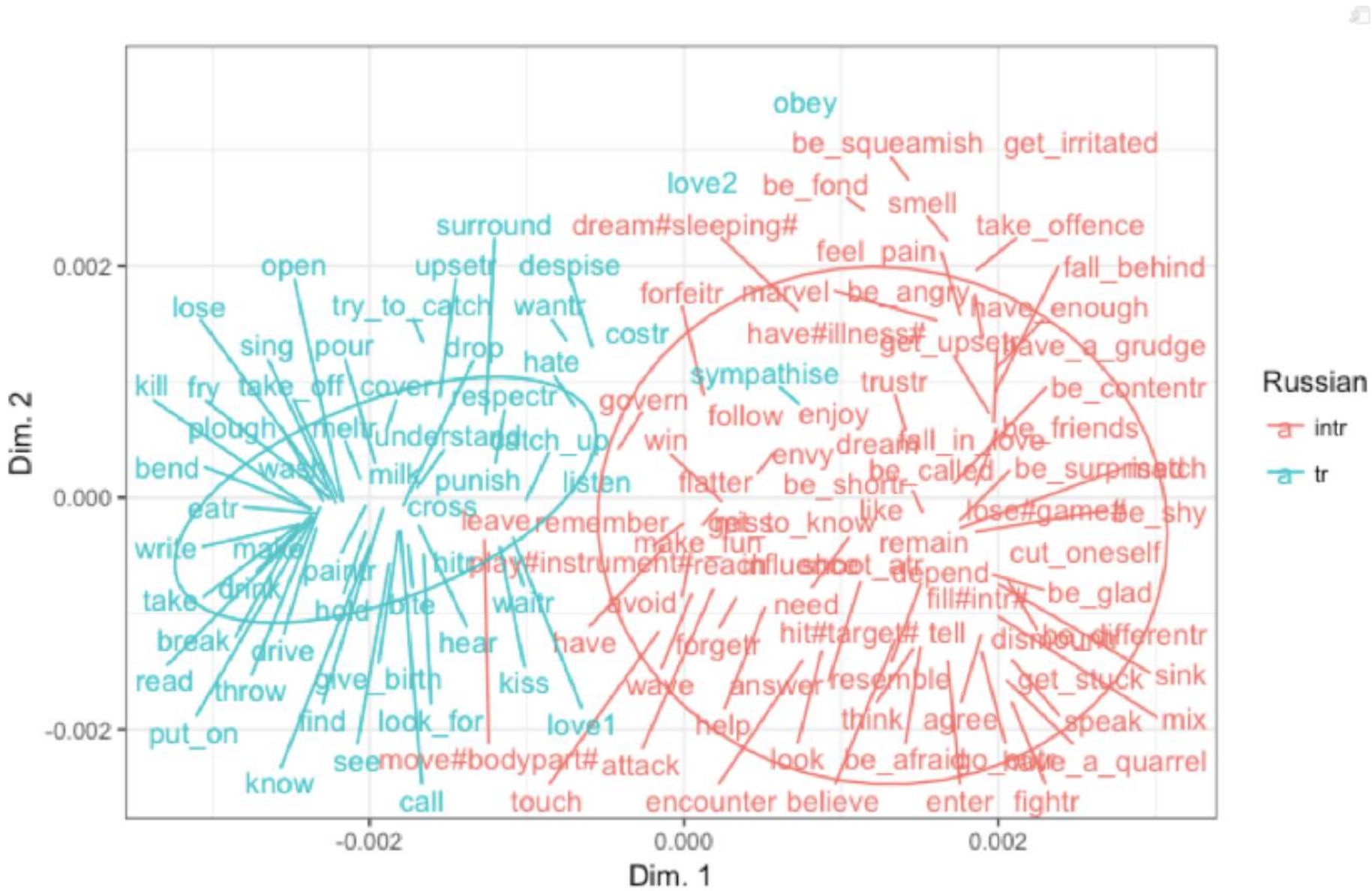

0.001
0.000
-0.001
-0.002
Dim. 2
-0.004
-0.002
0.000
0.002
Dim. 1
have_enough
govern
flatter
take_off
understand
throw
milk
punish
listen
be_called
drink
sing
find
catch_up
miss
be_shortr
write
eatr
need
take
put_on
hear
remember
shoot_atr
remain
break
read
call
kiss
love1
have
depend
make
know
look_for
waitr
reach
attack
forgetr
answer
hit#target#
hold
see
bite
drive
touch
help
believe
agree
look
go_outr
be_afraid
speak
encounter
enter
have_a_quarrel
fightr
Guarani
intr
tr



empty values… same is true for Tsaxur, Bagwalal and Khmer

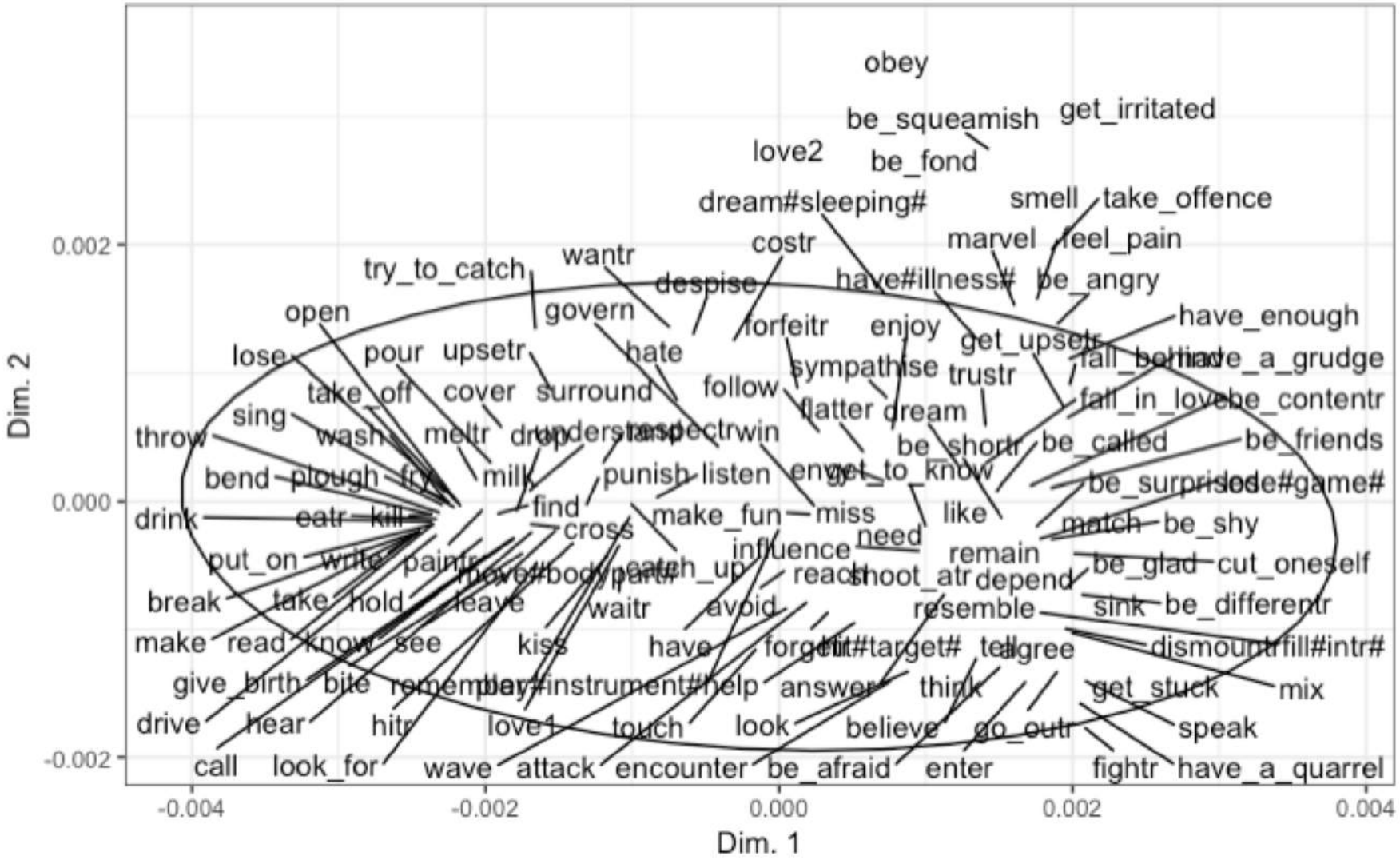
LET'S HAVE A LOOK AT OUTLIERS



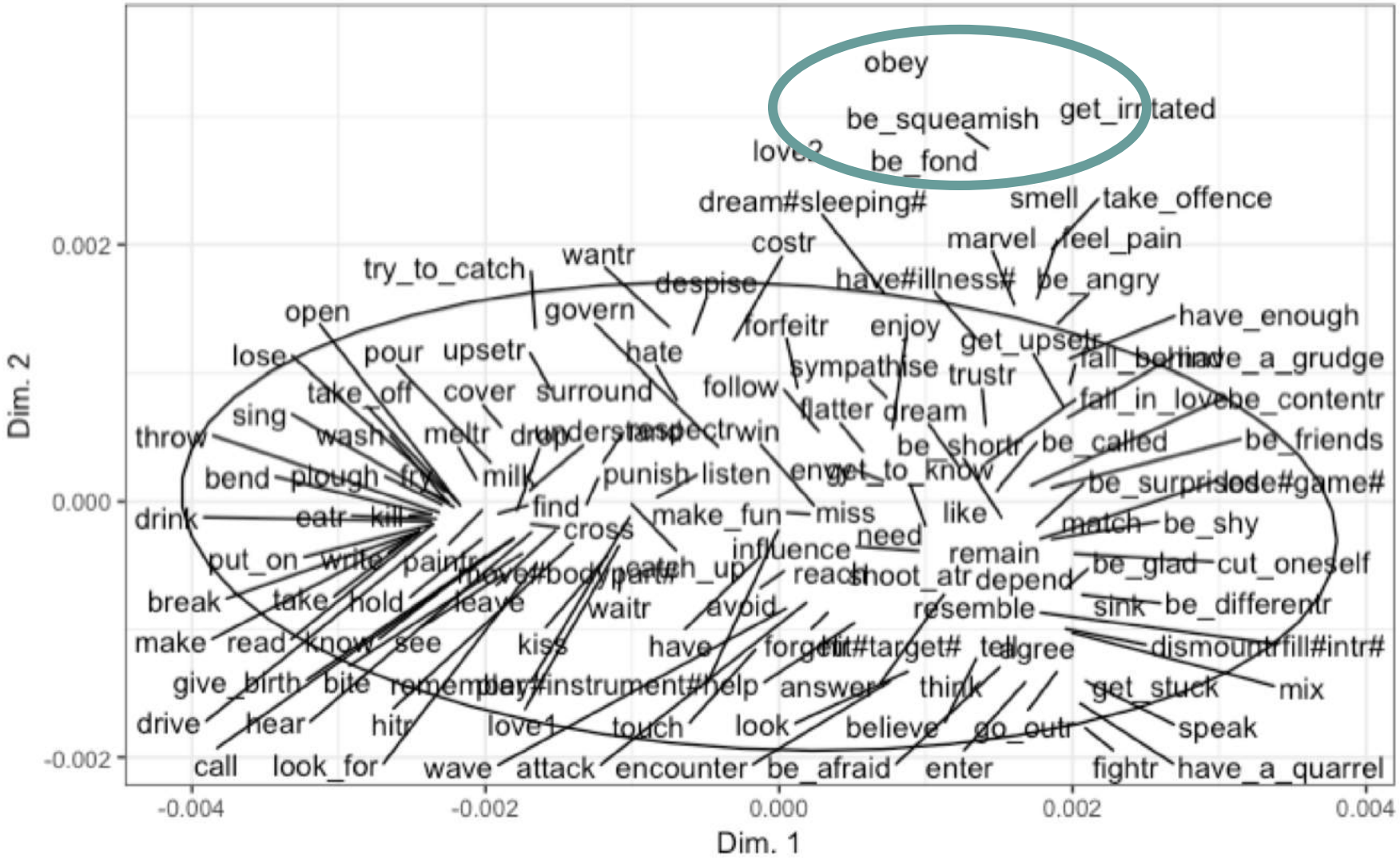


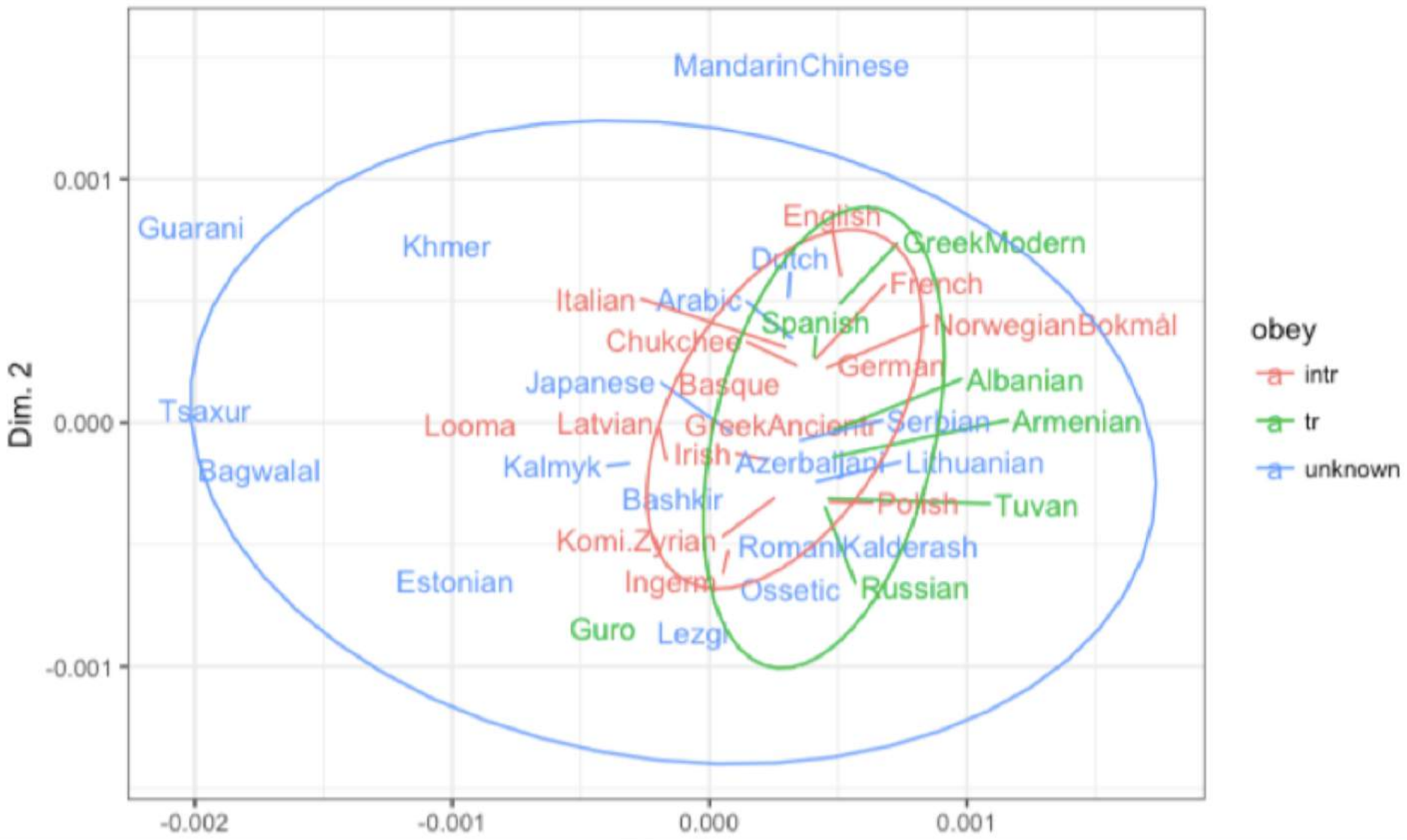
very intransitive

MCA FOR CONSTRUCTIONS



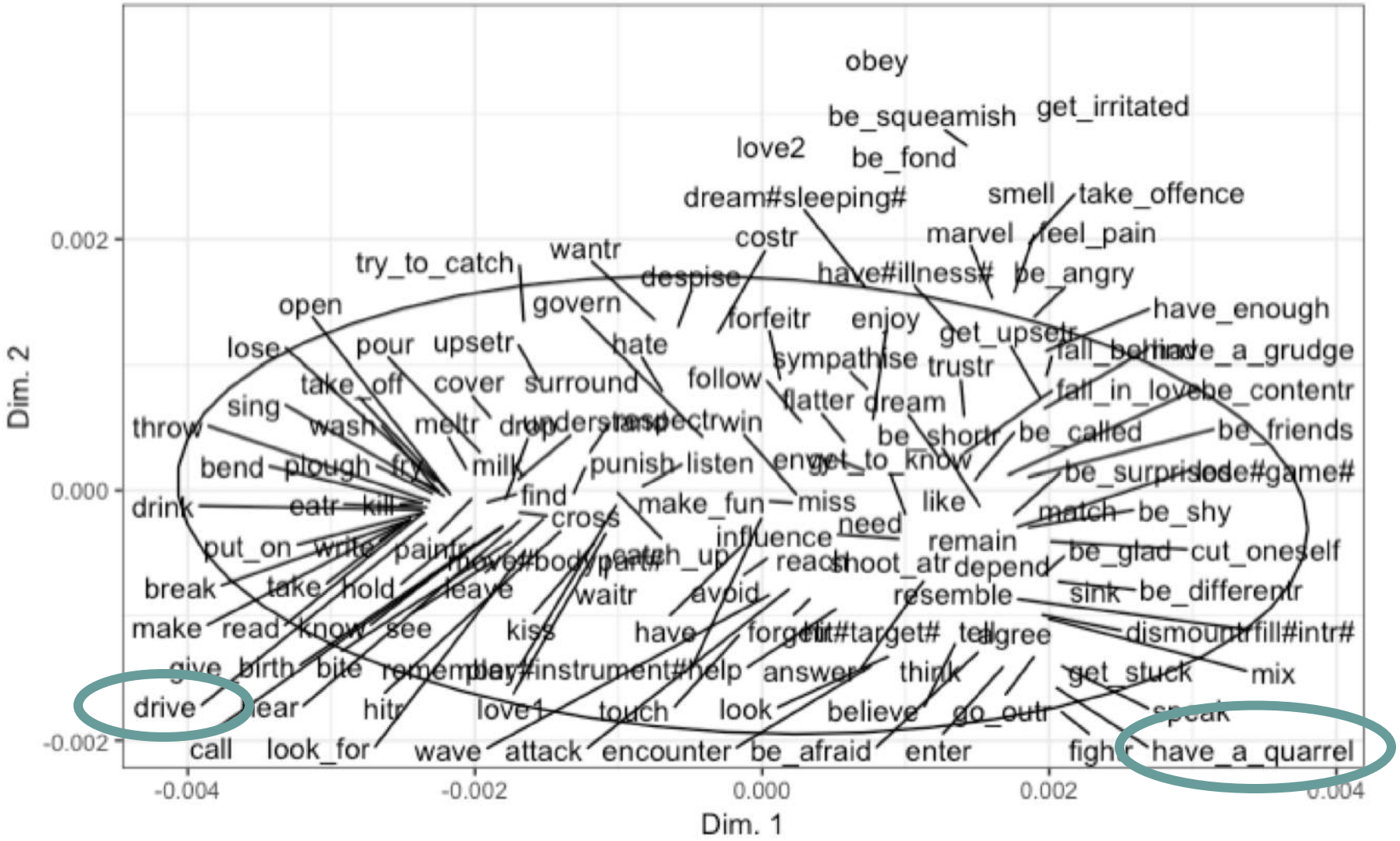
MCA FOR CONSTRUCTIONS

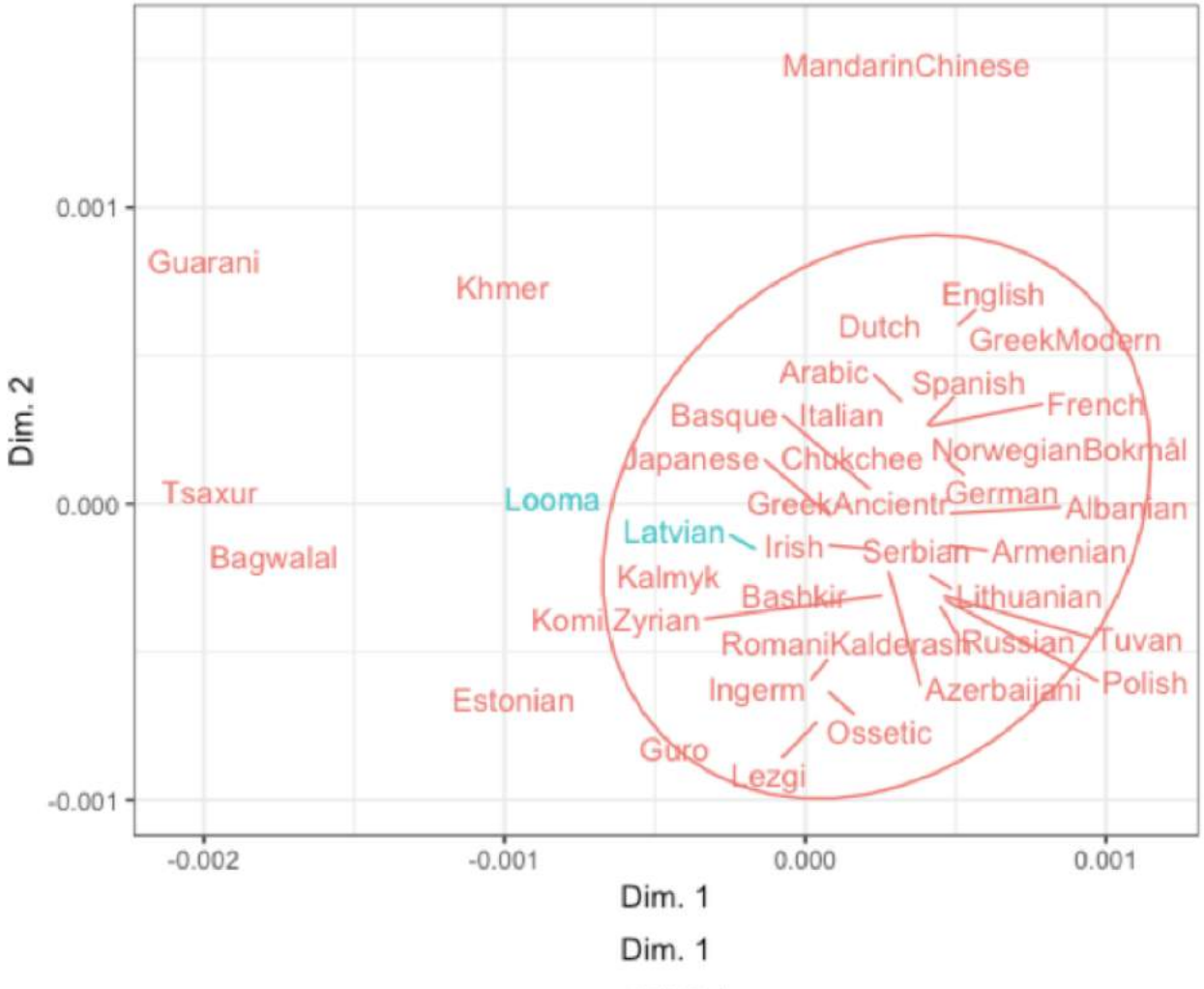




empty values…

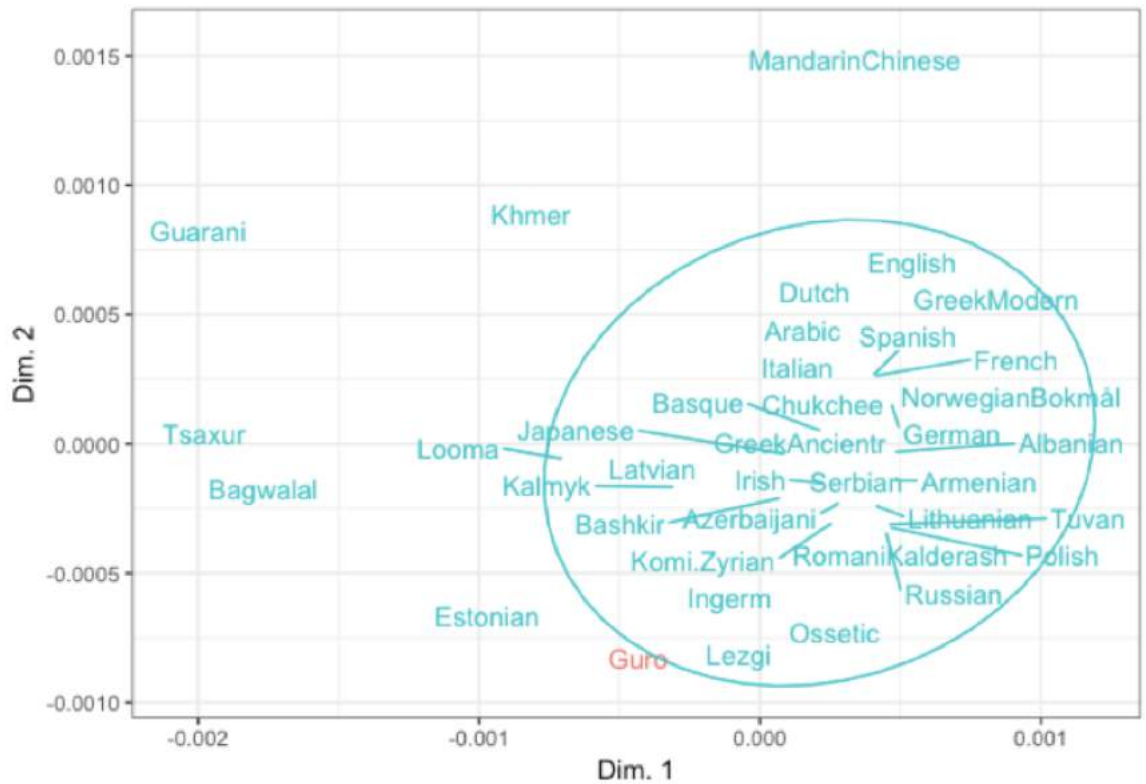
MCA FOR CONSTRUCTIONS





have_a_quarrel

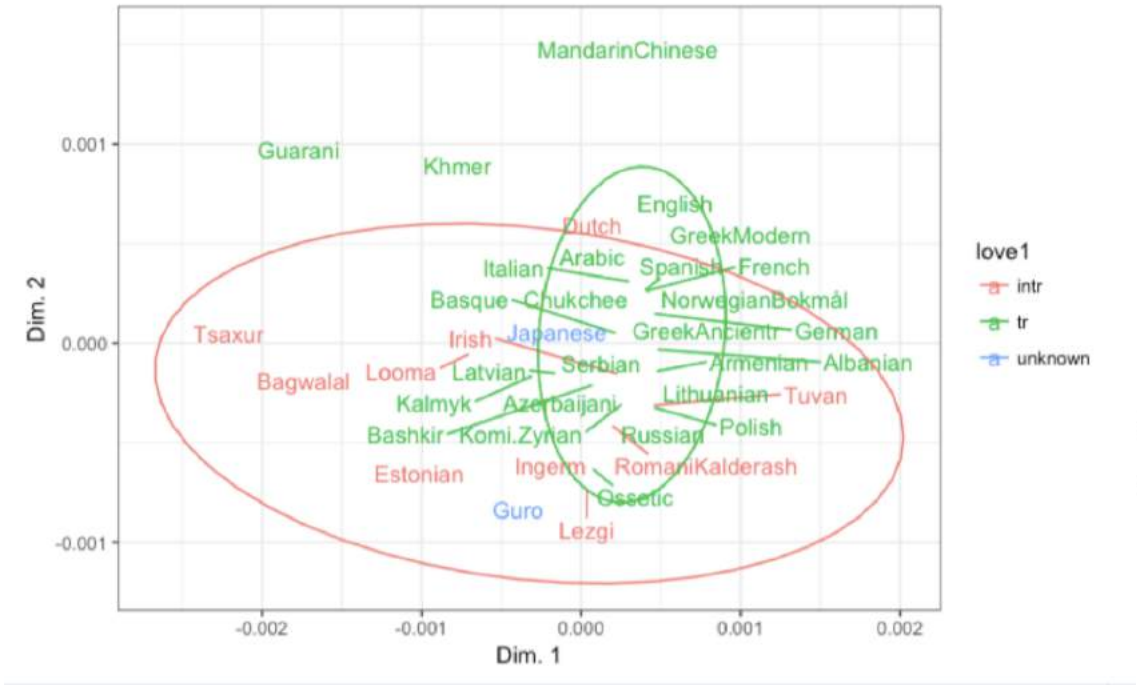
- intr
- unknown

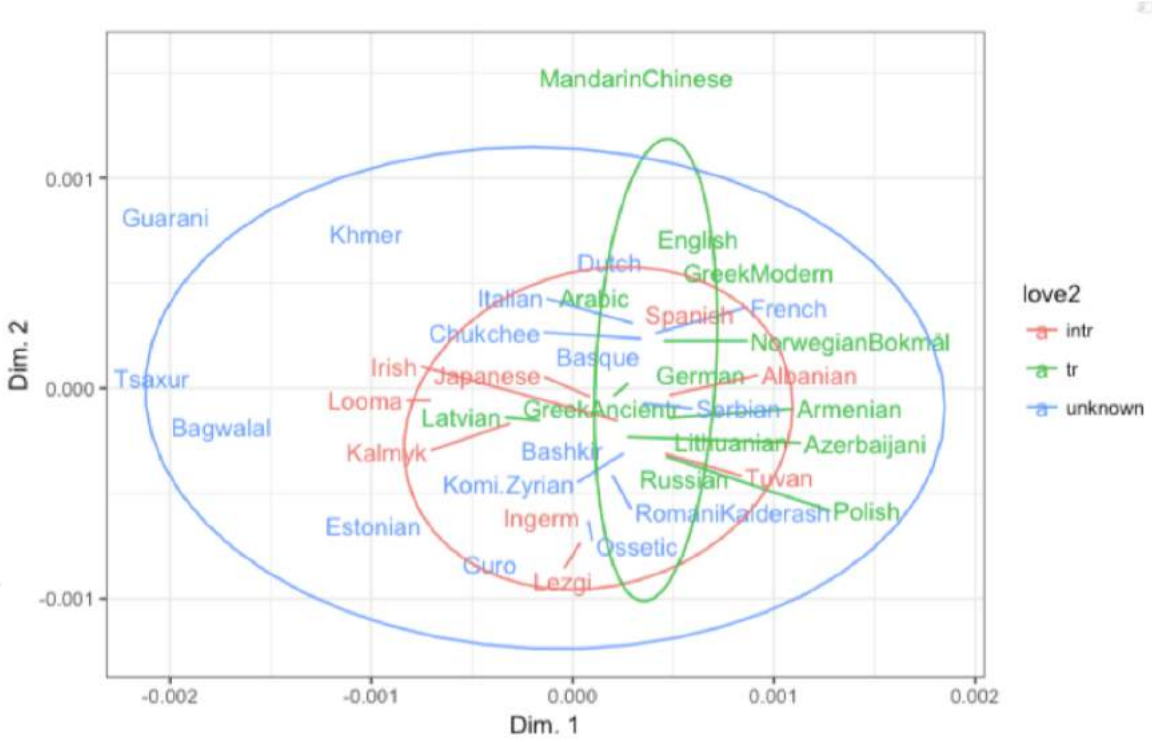


drive

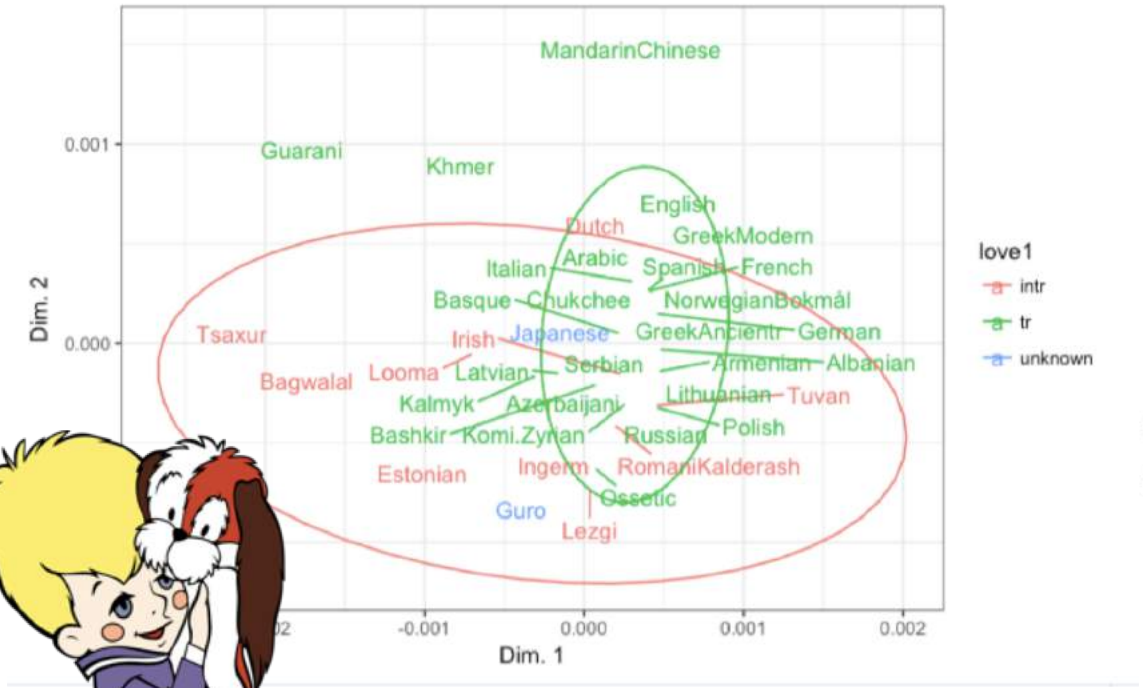
- intr
- tr

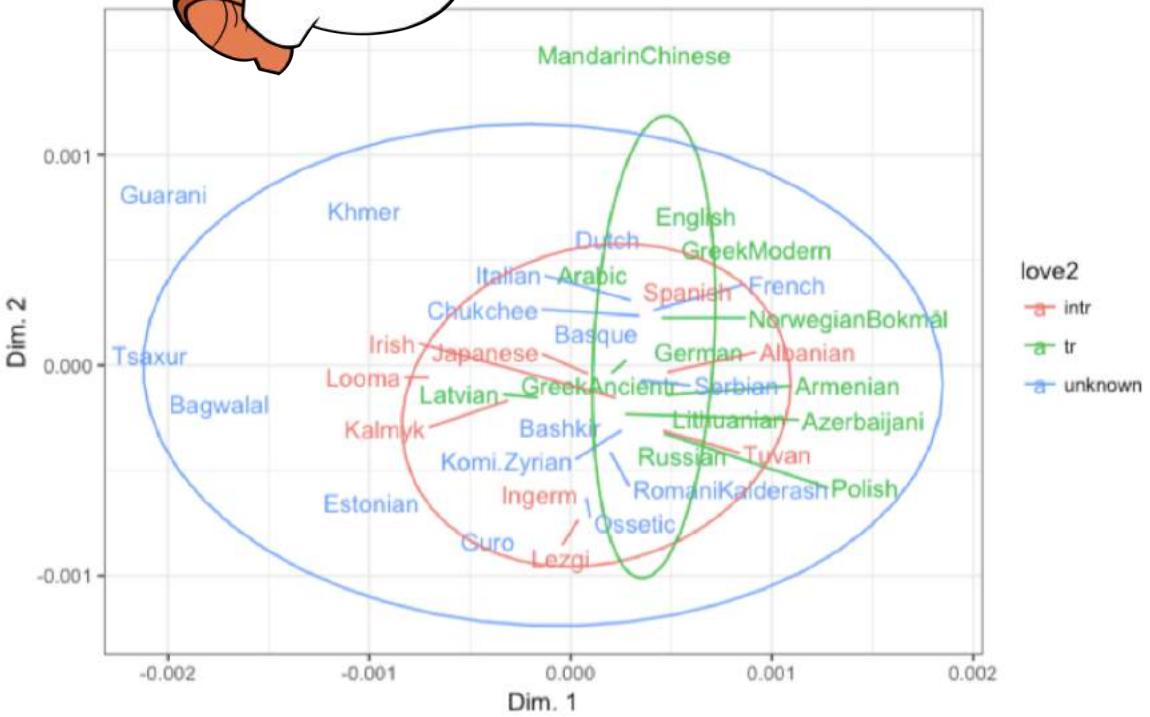
LOVE1 AND LOVE2

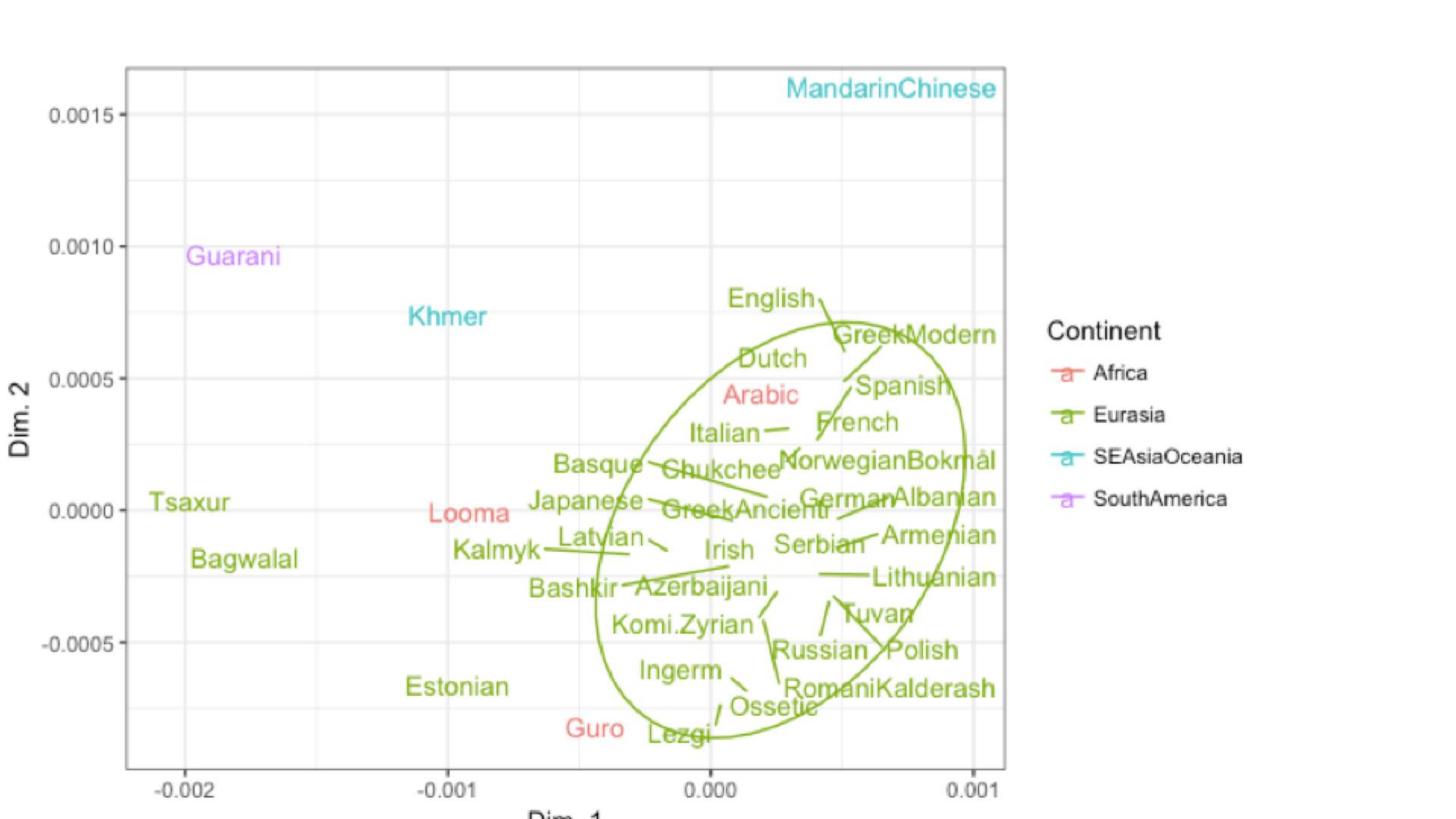


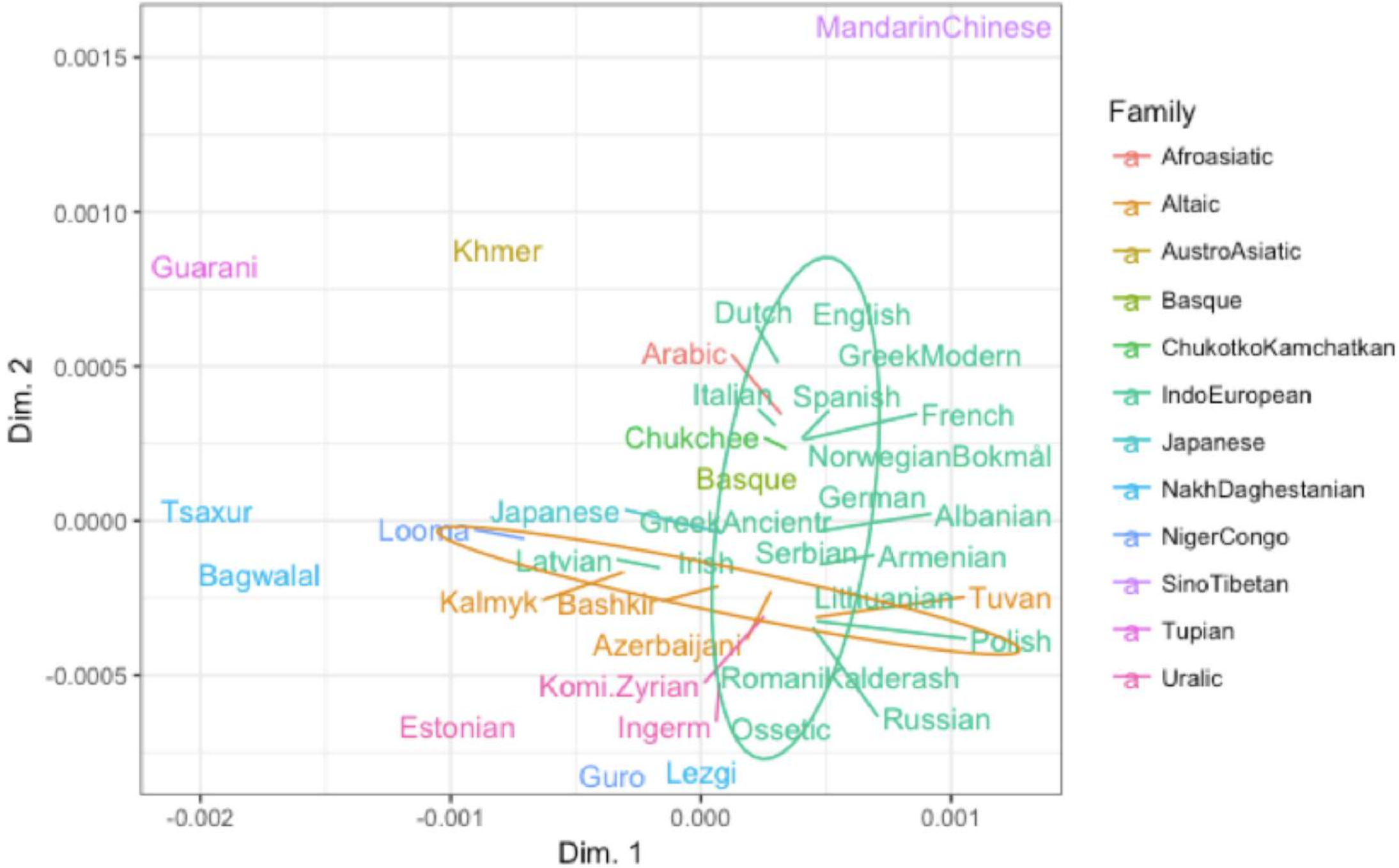


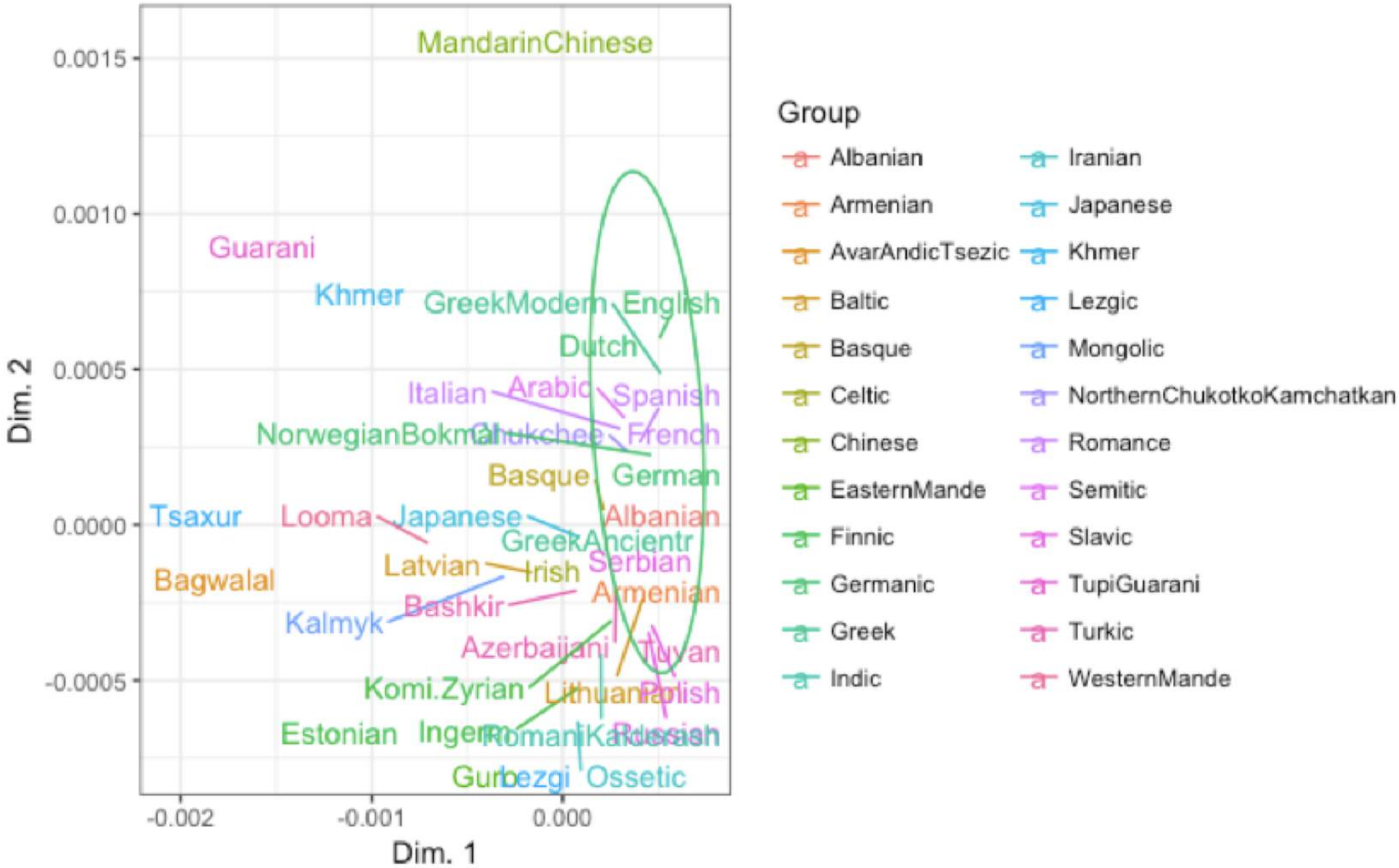
LOVE1 AND LOVE2






MandarinChinese
Guarani
Khmer
English
GreekModern
Dutch
Spanish
Arabic
French
Italian
NorwegianBokmål
Basque
Chukchee
Tsaxur
Looma
Japanese
GreekAncient
GermanAlbanian
Latvian
Armenian
Kalmyk
Irish
Serbian
Bagwalal
Lithuanian
Bashkir
Azerbaijani
Komi.Zyrian
Tuvan
Russian
Polish
Ingerm
Estonian
RomaniKalderash
Ossetic
Guro
Lezgi
0.0015
0.0010
0.0005
0.0000
-0.0005
-0.002
-0.001
0.000
0.001
Dim. 2
Dim. 1
Continent
Africa
Eurasia
SEAsiaOceania
SouthAmerica

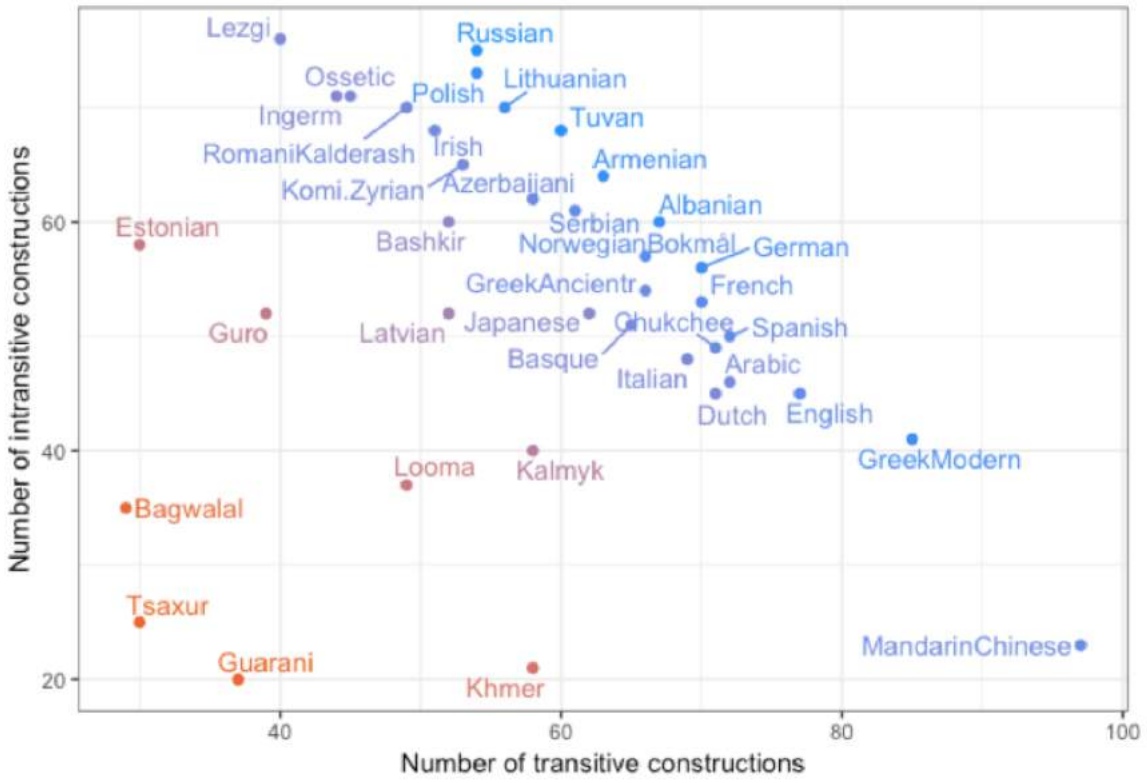

MandarinChinese
Khmer
Guarani
Dutch
English
Arabic
GreekModern
Italian
Spanish
French
Chukchee
NorwegianBokmål
Basque
German
Tsaxur
Japanese
Albanian
Loona
GreekAncientr
Serbian
Armenian
Latvian
Irish
Bagwalal
Kalmyk
Bashkir
Lithuanian
Tuvan
Polish
Azerbaijani
Komi.Zyrian
RomaniKalderash
Russian
Estonian
Ingerm
Ossetic
Guro
Lezgi
0.0015
0.0010
0.0005
0.0000
-0.0005
Dim. 2
-0.002
-0.001
0.000
0.001
Dim. 1
Family
Afroasiatic
Altaic
AustroAsiatic
Basque
ChukotkoKamchatkan
IndoEuropean
Japanese
NakhDaghestanian
NigerCongo
SinoTibetan
Tupian
Uralic


MandarinChinese
Guarani
Khmer
GreekModern
English
Dutch
Italian
Arabic
Spanish
NorwegianBokmal
Chukchee
French
Basque
German
Tsaxur
Looma
Japanese
Albanian
GreekAncientr
Latvian
Irish
Serbian
Bagwalal
Armenian
Bashkir
Kalmyk
Azerbaijani
Tuvan
Komi.Zyrian
Lithuanian
Polish
Estonian
Ingrian
Romani
Kabardian
Russian
Gurbezgi
Ossetic
0.0015
0.0010
0.0005
0.0000
-0.0005
Dim. 2
-0.002
-0.001
0.000
Dim. 1
Group
Albanian
Armenian
AvarAndicTsezic
Baltic
Basque
Celtic
Chinese
EasternMande
Finnic
Germanic
Greek
Indic
Iranian
Japanese
Khmer
Lezgic
Mongolic
NorthernChukotkoKamchatkan
Romance
Semitic
Slavic
TupiGuarani
Turkic
WesternMande

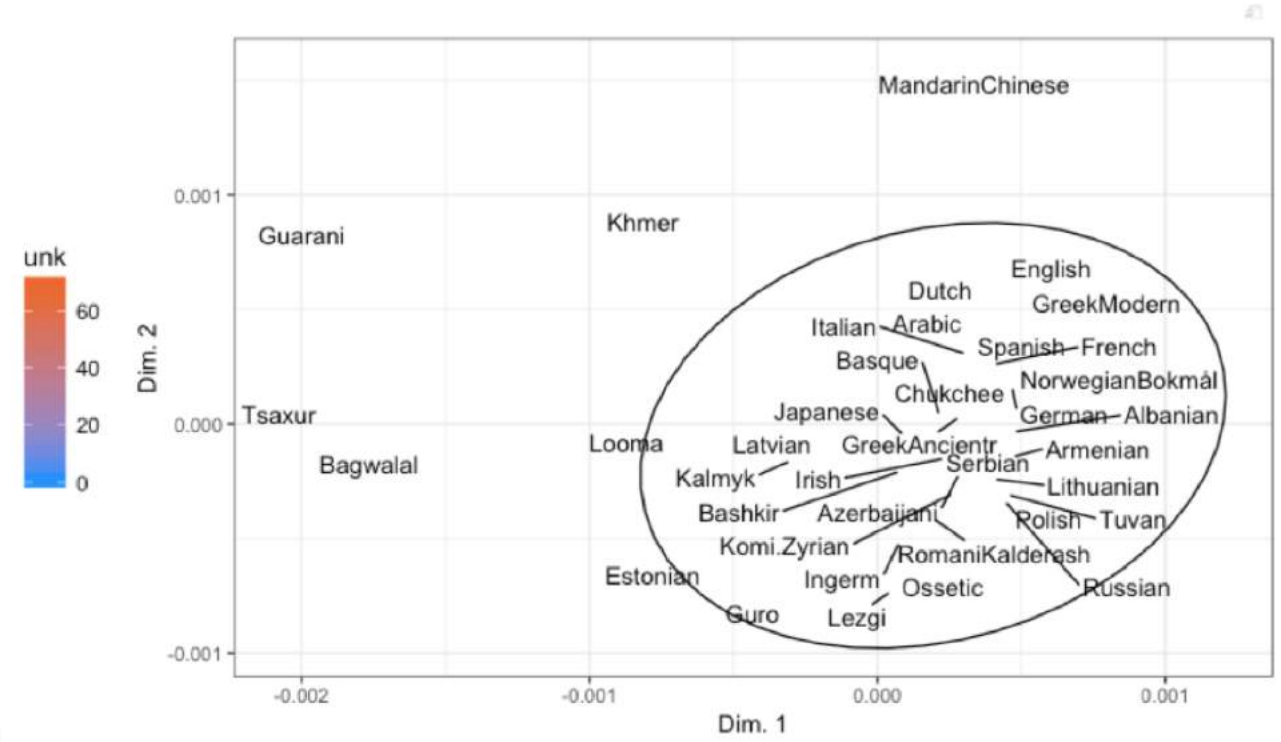
COMPARISON OF THE RESULTS PRODUCED BY DIFFERENT METHODS

LANGUAGES

Scatter plot

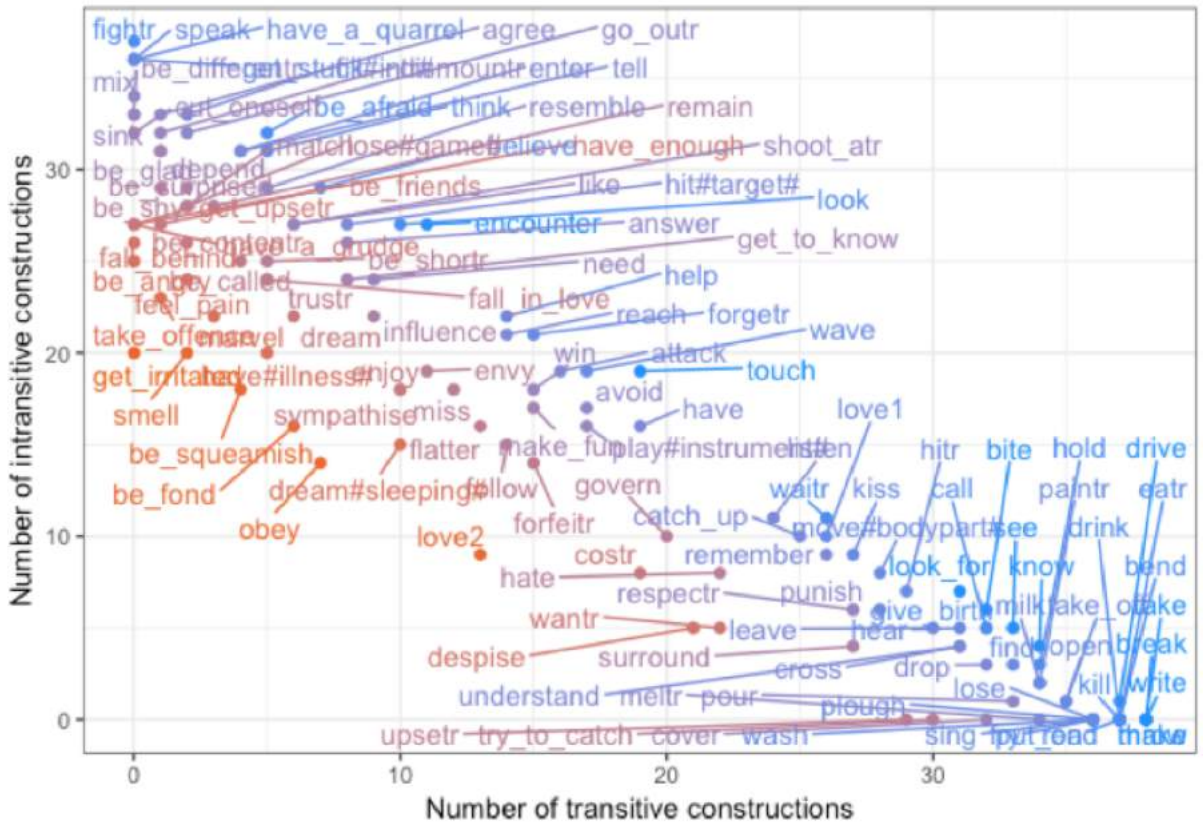


MCA

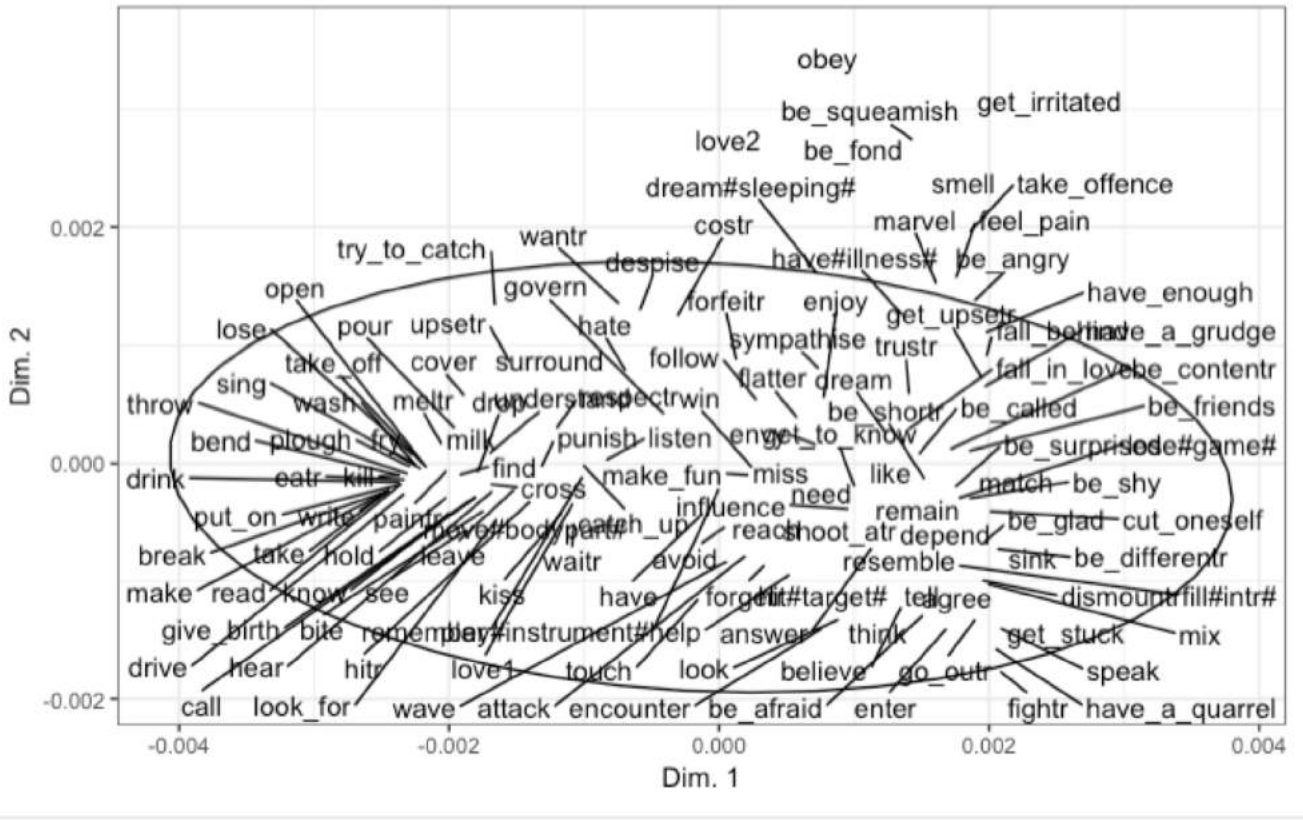


CONSTRUCTIONS

Scatter plot



MCA



THANK YOU FOR YOUR ATTENTION!

REFERENCES

Hopper, Paul J. and Thompson, Sandra A. 1980. Transitivity in grammar and discourse. *Language* 56: 251–299.

Say, Sergey. 2014. Bivalent verb classes in the languages of Europe. *Language Dynamics and Change* 4.1: 116-166.