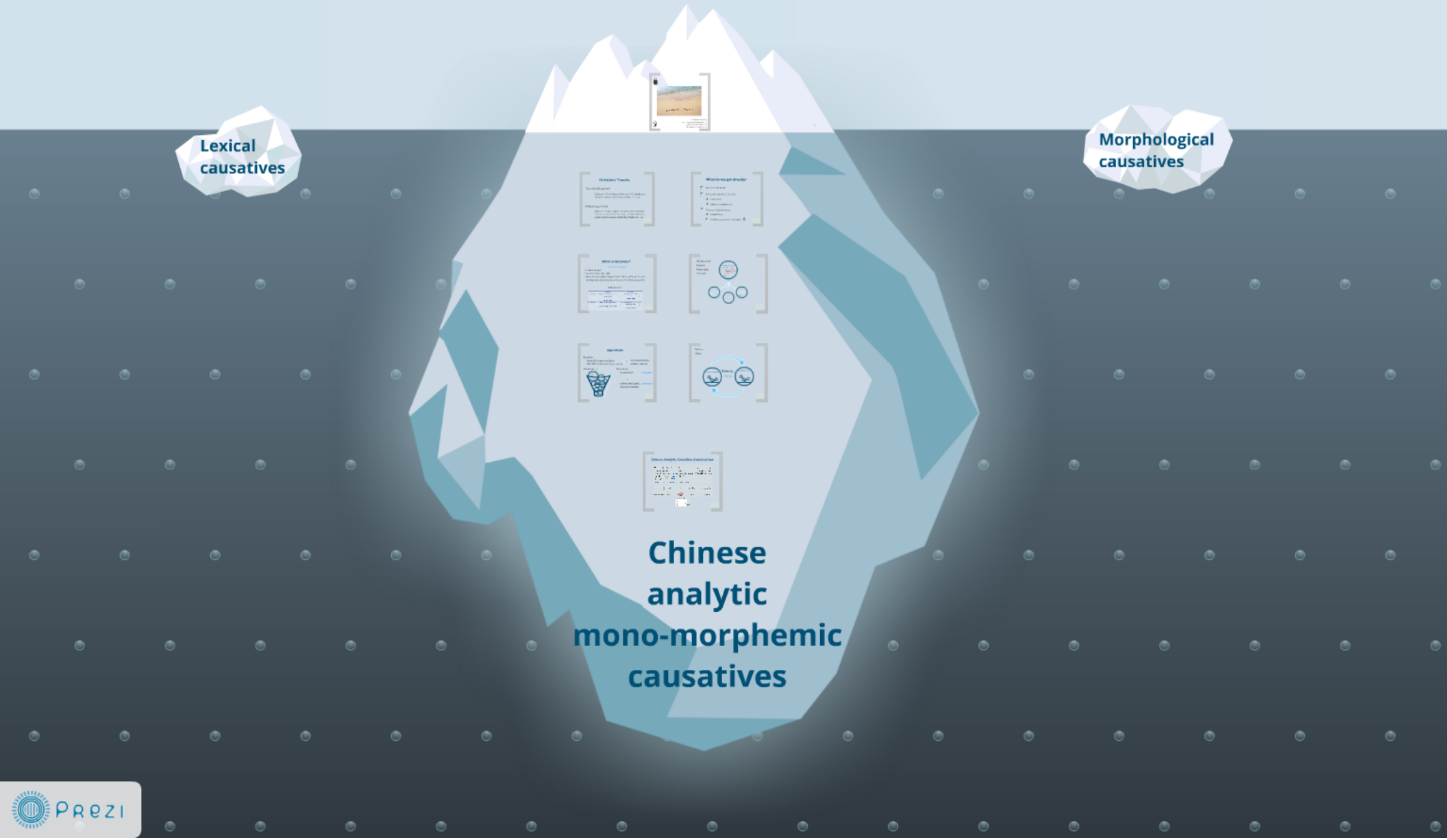
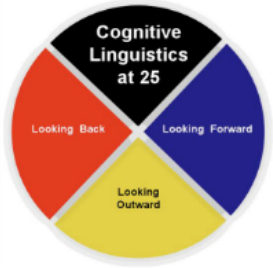
CAUSE cross Chronology of Chinese:
A corpus-based analysis of Chinese mono-morphemic causatives

Yanan Hu & Dirk Speelman

Research Unit QLVL
Department of Linguistics, Faculty of Arts
KU Leuven



Prezi



CAUSE cross Chronology of Chinese: A corpus-based analysis of Chinese mono-morphemic causatives

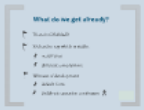
Yanan Hu & Dirk Speelman

Research Unit QLVL
Department of Linguistics, Faculty of Arts
KU Leuven

KU Leuven

Lexical causatives

Morphological causatives



Chinese analytic mono-morphemic causatives

Prezi

Chinese Analytic Causative Construction

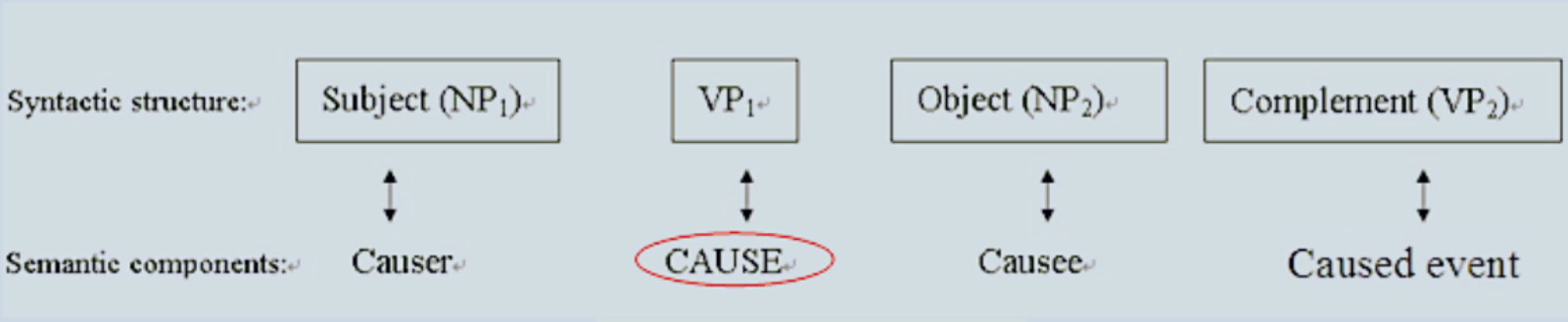
我	让	客	人	围	着	桌	子	坐	下。
Wǒ	ràng	kè	rén	wéi	zhe	zhuō	zi	zuò	xià

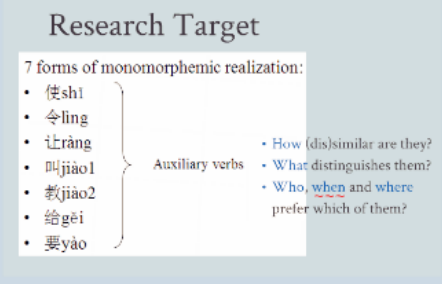
I CAUSE the guests surround (present tense marker) the table sit down

I asked the guests to sit around the table.

NP1 + VP1 + NP2 + VP2

Causer + CAUSE + Causee + Caused event





Prezi

Research Target

7 forms of monomorphemic realization:

- 使shǐ
- 令lìng
- 让ràng
- 叫jiào1
- 教jiào2
- 给gěi
- 要yào

Auxiliary verbs

- How (dis)similar are they?
- What distinguishes them?
- Who, when and where prefer which of them?

Prezi

Forefathers' Treasure

Theoretical Background:

Shibatani 1976, Verhagen & Kemmer 1997, Speelman & Geeraerts 2009, Geeraerts 2010, Levshina 2011, etc.

Methodological Tools:

Statistics for Corpus Linguistics (Speelman 1997), distinctive collexeme analysis (Gries & Stefanowitsch 2004), multinomial logistic regression analysis, motion chart (Hilpert 2011), etc.

Prezi

What to Excavate?

Research Questions:

- Is there change?
- Which is relatively stable?
- How do the so-called Chinese "doen" ("shi") and "laten" ("rang") develop, especially along the continuum of (in)direct causation?

↓

Geeraerts 2010

	synchrony	diachrony
semasiology	polysemous senses/usages of one causal auxiliary ↑ polysemy study	meaning change over time ↑ semantic change
onomasiology	nuances between near-synonyms ↑ synonymy study, variation study	lexical preference difference for one concept over time ↑ expression variation

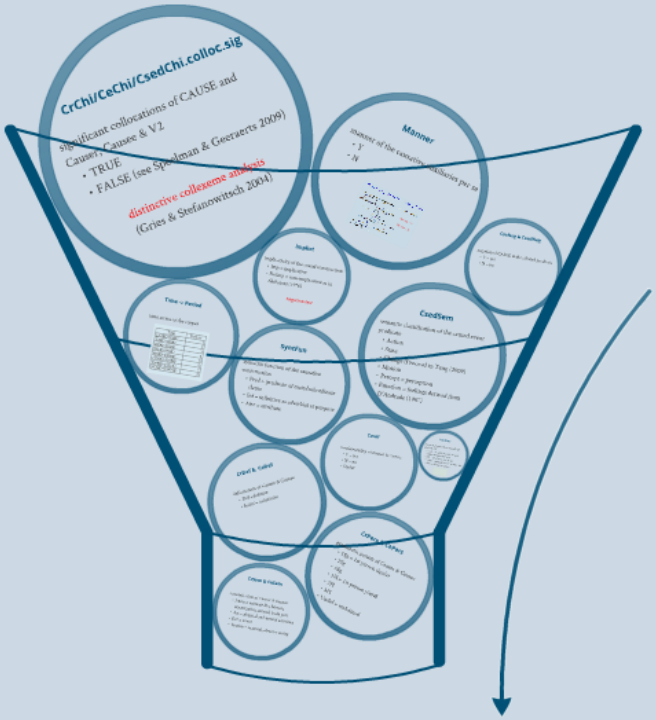
Prezi

Operation

Materials:

- Sheffield Corpus of Chinese → 2,531 observations
- The UCLA Chinese Corpus (1st ed) → 1100BC-2005AD

Predictors: 18

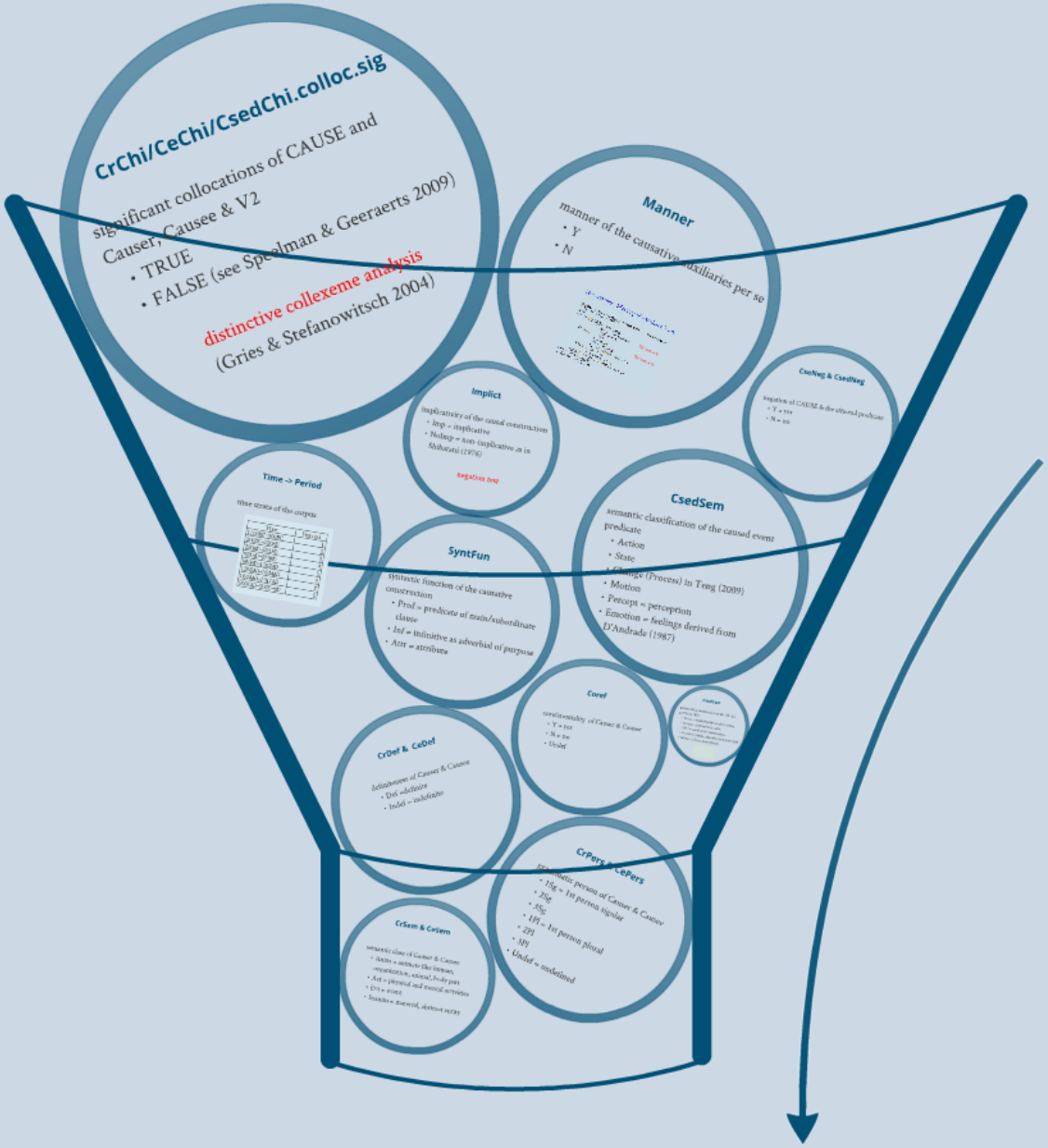


Procedures:

- motion chart (exploratory)
- multinomial logistic regression analysis (confirmatory)

Prezi

Predictors: 18



Proce

- m
- m
  re

Prezi

CrSem & CeSem

semantic class of Causer & Causee

- Anim = animate like human, organization, animal, body part
- Act = physical and mental activities
- Evt = event
- Inanim = material, abstract entity

Prezi

f purpose

Coref

coreferentiality of Causer & Causee

- Y = yes
- N = no
- Undef

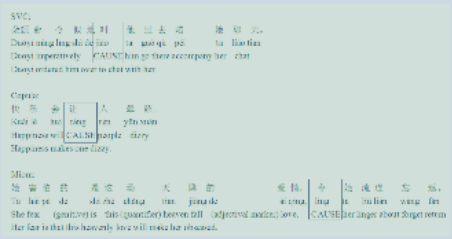
grammatical c
predicate (V2)
- Trans = tr
- Intrans = i
- SVC = ser
- Copula = c
- Idiom = id

Prezi

CsedCstr

grammatical construction of the effected predicate (V2)

- Trans = transitive/ditransitive verbs
- Intrans = intransitive verbs
- SVC = serial verb construction
- Copula = copula, adjective, past participle
- Idiom = idiom, noun phrase



SVC:

朵颐 命 令 似 地 叫 他 过 去 陪 她 聊 天。

Duǒyí mìng lìng shì de jiào tā guò qù péi tā liáo tiān

Duoyi imperatively CAUSE him go there accompany her chat

Duoyi ordered him over to chat with her.

Copula:

快 乐 会 让 人 晕 眩。

Kuài lè huì ràng rén yūn xuàn

Happiness will CAUSE people dizzy

Happiness makes one dizzy.

Idiom:

她 害 怕 的 是 这 场 天 降 的 爱 情， 令 她 流 连 忘 返。

Tā hài pà de shì zhè chǎng tiān jiàng de ài qíng, lìng tā liú lián wàng fǎn

She fear (genitive) is this (quantifier) heaven fall (adjectival marker) love, CAUSE her linger about forget return

Her fear is that this heavenly love will make her obsessed.

Prezi

CsedSem

semantic classification of the caused event predicate

- Action
- State
- Change (Process) in Teng (2009)
- Motion
- Percept = perception
- Emotion = feelings derived from D'Andrade (1987)

nate

Prezi

Time -> Period

time strata of the corpus

Time	Period
1100BC-206BC	1
206BC-220AD	2
220AD-581AD	3
581AD-979AD	4
860AD-1368AD	5
1368AD-1644AD	6
1644AD-1911AD	7
2000AD-2005-AD	8

synta

cons

- P
- I
- A

Prezi

CrChi/CeChi/CsedChi.colloc.sig

significant collocations of CAUSE and Causer, Causee & V2

- TRUE
- FALSE (see Speelman & Geeraerts 2009)

distinctive collexeme analysis (Gries & Stefanowitsch 2004)

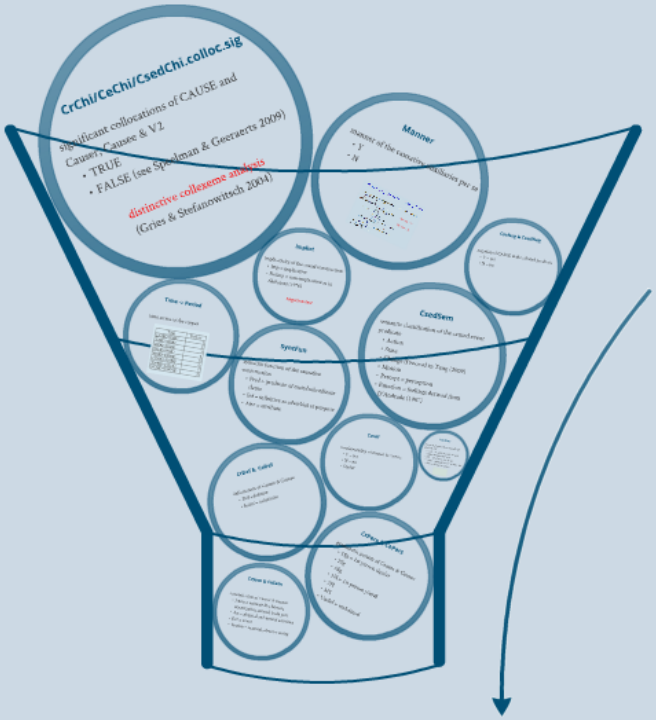
Prezi

Operation

Materials:

- Sheffield Corpus of Chinese → 2,531 observations
- The UCLA Chinese Corpus (1st ed) → 1100BC-2005AD

Predictors: 18

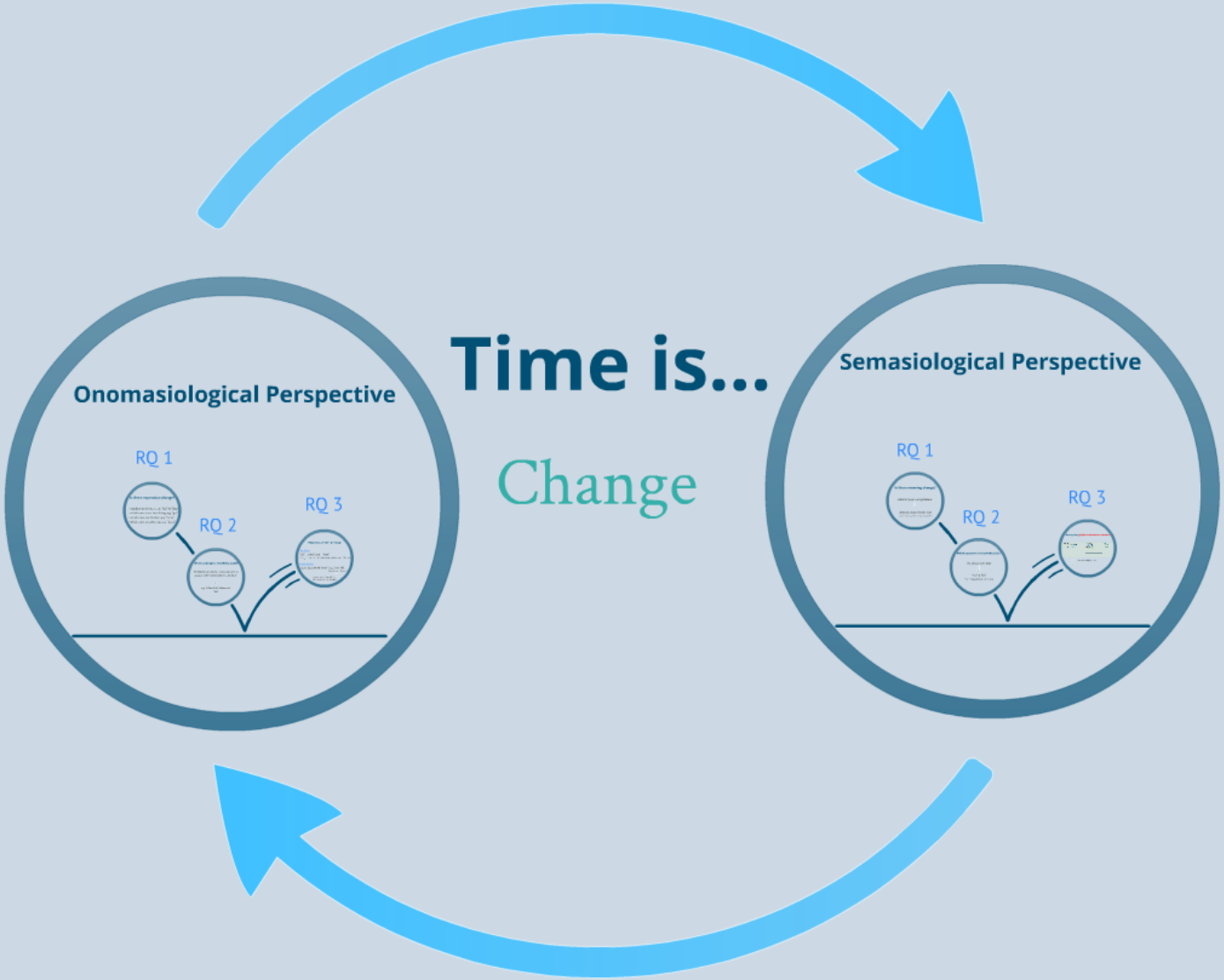


Procedures:

- motion chart (exploratory)
- multinomial logistic regression analysis (confirmatory)

Prezi

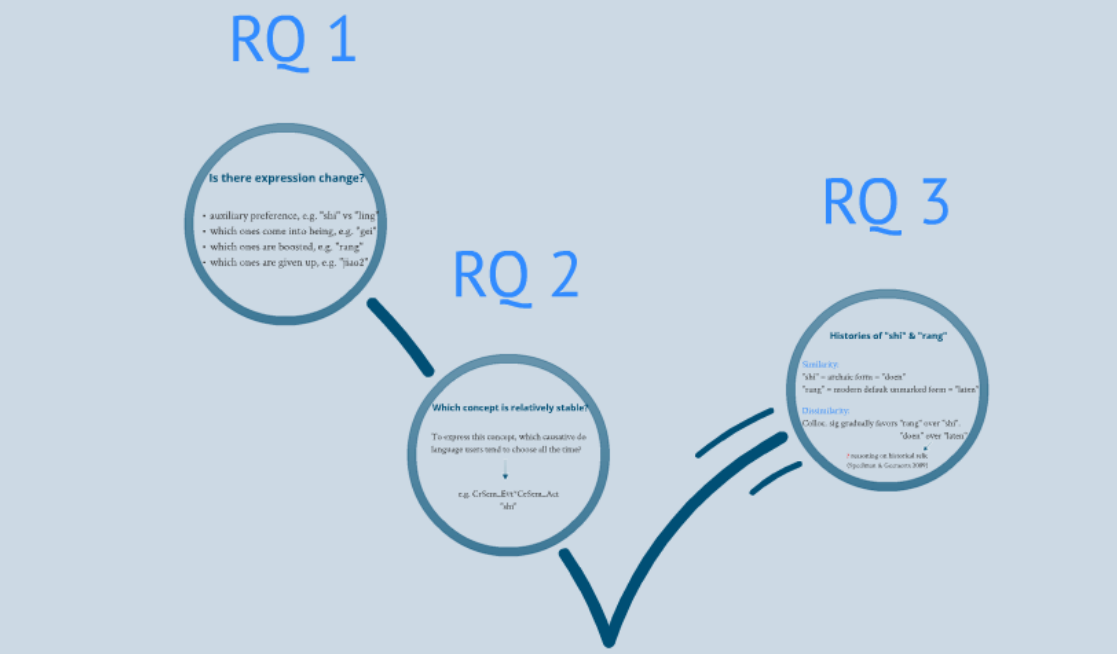
Motion Chart



Prezi

Tir

Onomasiological Perspective



Prezi

Is there expression change?

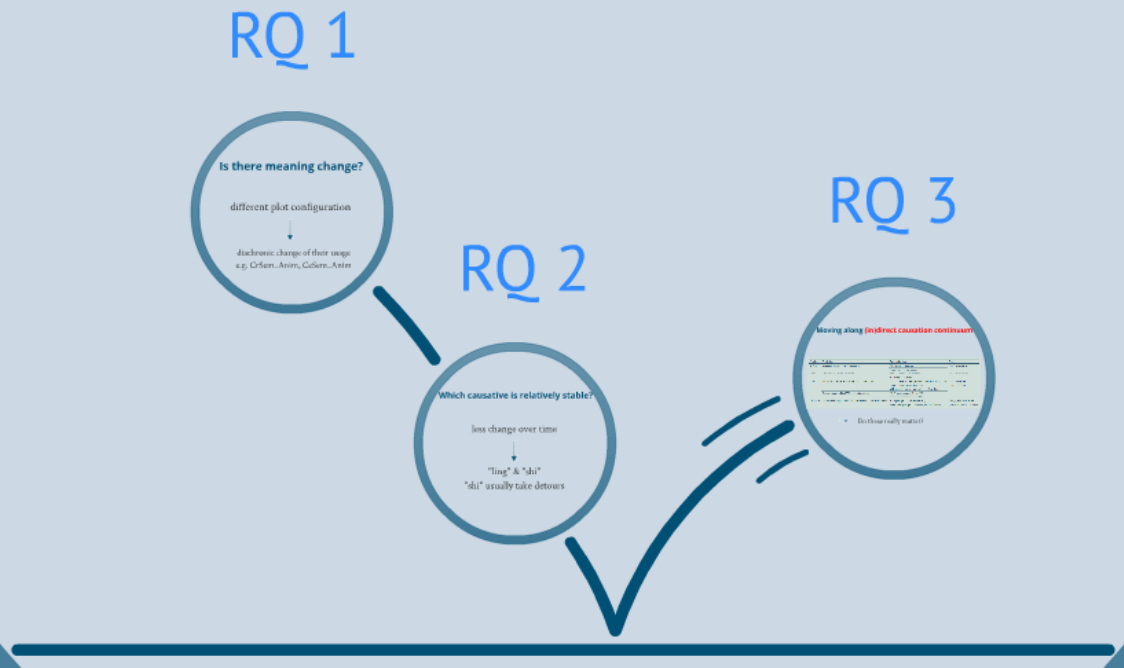
- auxiliary preference, e.g. "shi" vs "ling"
- which ones come into being, e.g. "gei"
- which ones are boosted, e.g. "rang"
- which ones are given up, e.g. "jiao2"

Prezi

s...

e

Semasiological Perspective



Prezi

Is there meaning change?

different plot configuration

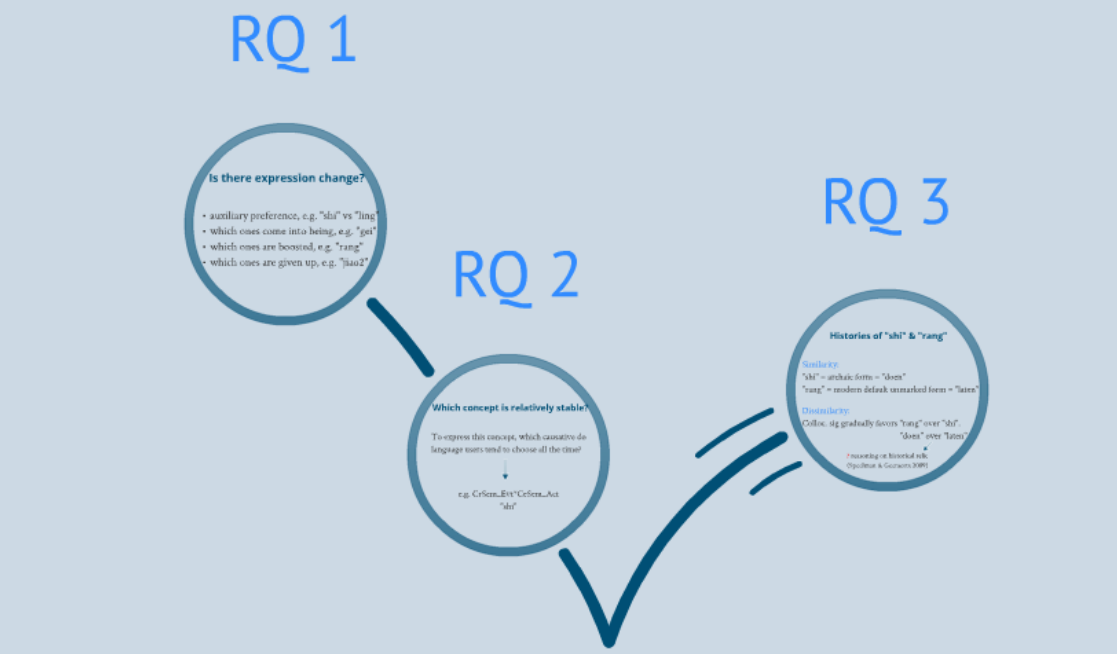
diachronic change of their usage

e.g. CrSem_Anim, CeSem_Anim

Prezi

Tir

Onomasiological Perspective



Prezi

Which concept is relatively stable?

To express this concept, which causative do language users tend to choose all the time?

e.g. CrSem_Evt*CeSem_Act

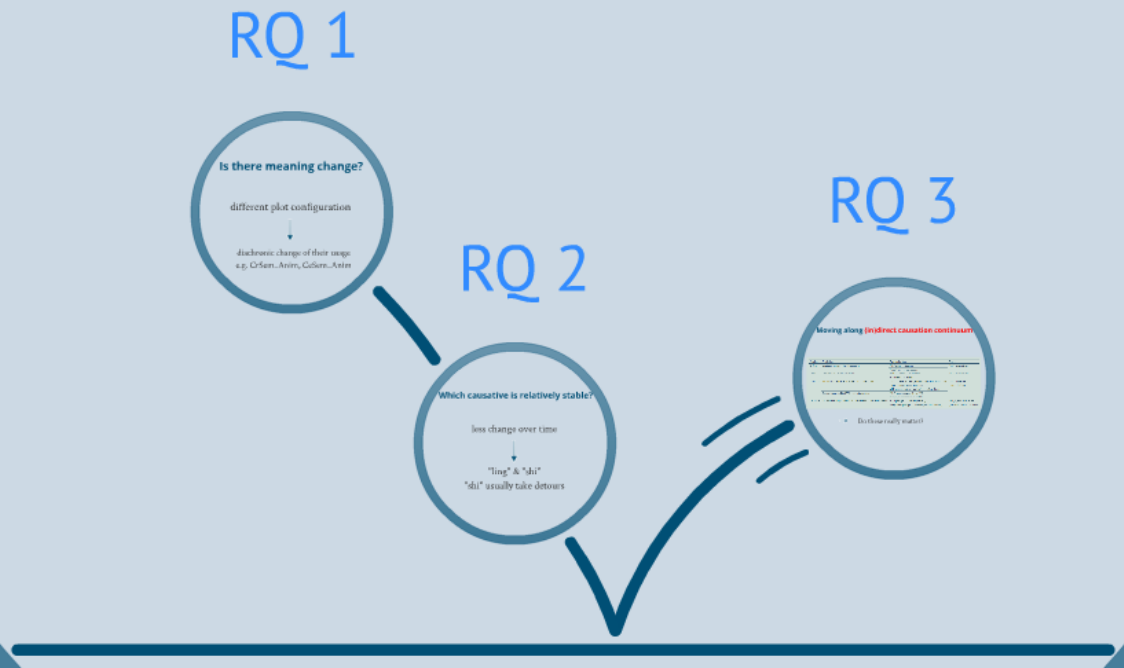
"shi"

Prezi

s...

e

Semasiological Perspective



Prezi

Which causative is relatively stable?

less change over time

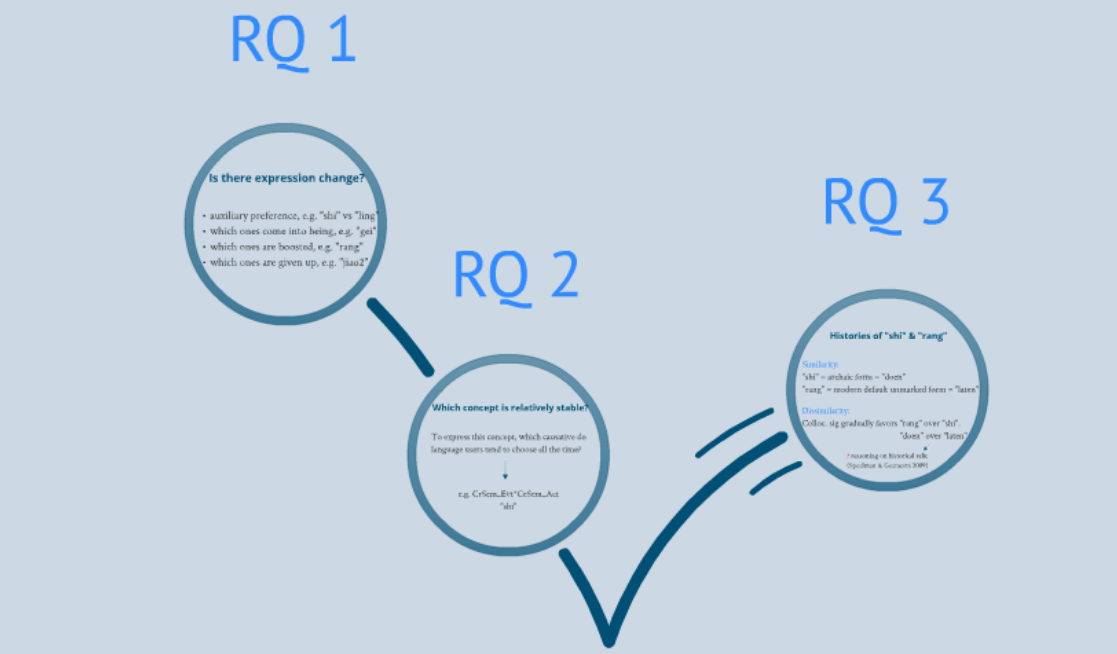
"ling" & "shi"

"shi" usually take detours

Prezi

Tir

Onomasiological Perspective



Prezi

Histories of "shi" & "rang"

Similarity:

"shi" = archaic form = "doen"

"rang" = modern default unmarked form = "laten"

Dissimilarity:

Colloc. sig gradually favors "rang" over "shi".

"doen" over "laten"

? reasoning on historical relic

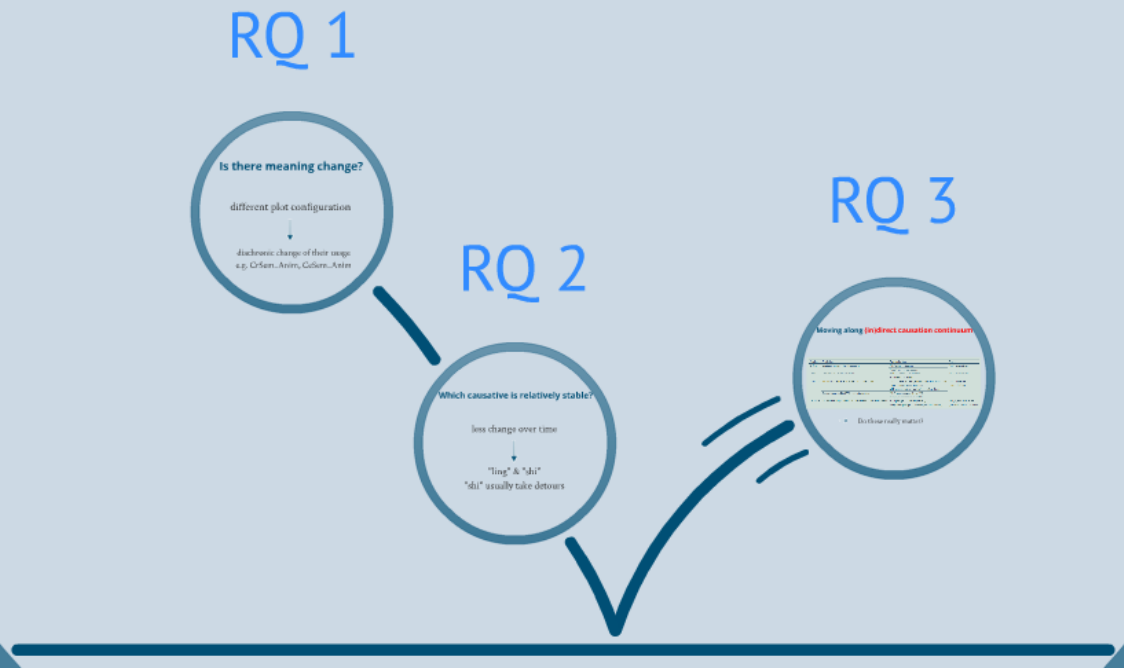
(Speelman & Geeraerts 2009)

Prezi

s...

e

Semasiological Perspective



Prezi

Moving along (in)direct causation continuum

Predictor	Prediction	Developing trace	Endpoint
CrSem	Inanimate causer = direct causation	"shi" anim → inanim	"shi", more direct
		"rang" too, but lags behind	
Coref	Coreference = direct causation	both non-coref → more coref "shi" coref > "rang"	"shi", more direct
CsedCstr	Intransitive (compared to transitive) = direct caustion	"shi", smaller percentage to take both trans and intrans	"shi", more direct; "rang", indirect
		"rang" intransitive → transitive a bit more transitive than "shi" in modern times	
	Copula (compared to SVC) = direct causation	"shi", more copulas, fewer SVC	
		"rang", more copulas, more SVC	
CsedSem	Perception verb (compared to emotion verb) = direct causation	"shi" percept → emotion (2 times)	"rang", a bit more direct (towards different direction)
		"rang" only percept → emotion (not up to 2 times)	

→ Do these really matter?

Prezi

Moving along (in)direct causation continuum

Predictor	Prediction	Developing trace	Endpoint
CrSem	Inanimate causer = direct causation	"shi" anim → inanim	"shi", more direct
		"rang" too, but lags behind	
Coref	Coreference = direct causation	both non-coref → more coref "shi" coref > "rang"	"shi", more direct
CsedCstr	Intransitive (compared to transitive) = direct caustion	"shi", smaller percentage to take both trans and intrans	"shi", more direct; "rang", indirect
		"rang" intransitive → transitive a bit more transitive than "shi" in modern times	
	Copula (compared to SVC) = direct causation	"shi", more copulas, fewer SVC	
		"rang", more copulas, more SVC	
CsedSem	Perception verb (compared to emotion verb) = direct causation	"shi" percept → emotion (2 times)	"rang", a bit more direct (towards different direction)
		"rang" only percept → emotion (not up to 2 times)	

→ Do these really matter?

Prezi

Multinomial Logistic Regression Analysis

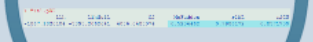
Multinomial Regression Model (ref="rang")

```
d$Causatives=relevel(d$Causatives,ref="rang")

fit1=multinom(Causatives~CrSem + CrDef + CrPers +
                CeSem + CeDef + CePers + Coref +
                Manner + CseNeg + CsedCstr + CsedSem +
                CsedNeg + Implict + SyntFun + Time +
                CrChi.colloc.sig +CeChi.colloc.sig+
                CsedChi.colloc.sig,
              data =d,
              maxit=1000)
```

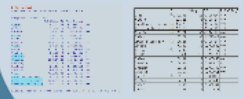
Model: Good enough?

pseudo R square

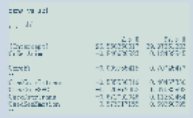


Predictors: Important?

Anova



Levels: Significant?



Prezi

Multinomial Regression Model (ref="rang")

```
d$Causatives=relevel(d$Causatives,ref="rang")

fit1=multinom(Causatives~CrSem + CrDef + CrPers +
                    CeSem + CeDef + CePers + Coref +
                    Manner + CseNeg + CsedCstr + CsedSem +
                    CsedNeg + Implict + SyntFun + Time +
                    CrChi.colloc.sig +CeChi.colloc.sig+
                    CsedChi.colloc.sig,
              data =d,
              maxit=1000)
```

Prezi

Model: Good enough?

pseudo R square

```
> fit1.pR2
          llh       llhNull            G2      McFadden          r2ML          r2CU
-1867.1820104 -3895.2062641  4056.0485074     0.5206462     0.7986172     0.8371700
```

Prezi

Predictors: Important?

Anova

```
> Anova(fit1)
Analysis of Deviance Table (Type II tests)

Response: Causatives
                   LR Chisq Df Pr(>Chisq)
CrSem                62.42 18  8.258e-07 ***
CrDef                11.43  6  0.0759486 .
CrPers               70.32 36  0.0005369 ***
CeSem                39.49 18  0.0024445 **
CeDef                10.57  6  0.1025079
CePers              103.79 36  1.717e-08 ***
Coref                20.15 12  0.0643229 .
Manner              417.45  6  < 2.2e-16 ***
CseNeg               14.11  6  0.0284492 *
CsedCstr             39.46 24  0.0244060 *
CsedSem             186.07 30  < 2.2e-16 ***
CsedNeg              25.05  6  0.0003337 ***
Implict              22.58  6  0.0009511 ***
SyntFun              59.10 12  3.287e-08 ***
Time               1503.53 42  < 2.2e-16 ***
CrChi.colloc.sig     94.26  6  < 2.2e-16 ***
CeChi.colloc.sig    206.99  6  < 2.2e-16 ***
CsedChi.colloc.sig  242.73  6  < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

	LR	Chisq	Df	Pr(>Chisq)
Time	1503.53	42	<2.2e-16	***
Manner	417.45	6	<2.2e-16	***
CsedChi.colloc.sig	242.73	6	<2.2e-16	***
CeChi.colloc.sig	206.99	6	<2.2e-16	***
CsedSem	186.07	30	<2.2e-16	***
CePers	103.79	36	1.72E-08	***
CrChi.colloc.sig	94.26	6	<2.2e-16	***
CrPers	70.32	36	0.000537	***
CrSem	62.42	18	8.26E-07	***
SyntFun	59.1	12	3.29E-08	***
CeSem	39.49	18	0.002445	**
CsedCstr	39.46	24	0.024406	*
CsedNeg	25.05	6	0.000334	***
Implict	22.58	6	0.000951	***
Coref	20.15	12	0.064323	.
CseNeg	14.11	6	0.028449	*
CrDef	11.43	6	0.075949	.
CeDef	10.57	6	0.102508	

Prezi

```
> Anova(fit1)
Analysis of Deviance Table (Type II tests)

Response: Causatives
                    LR Chisq Df Pr(>Chisq)
CrSem                 62.42 18  8.258e-07 ***
CrDef                 11.43  6  0.0759486 .
CrPers                70.32 36  0.0005369 ***
CeSem                 39.49 18  0.0024445 **
CeDef                 10.57  6  0.1025079
CePers               103.79 36  1.717e-08 ***
Coref                 20.15 12  0.0643229 .
Manner               417.45  6  < 2.2e-16 ***
CseNeg                14.11  6  0.0284492 *
CsedCstr              39.46 24  0.0244060 *
CsedSem              186.07 30  < 2.2e-16 ***
CsedNeg               25.05  6  0.0003337 ***
Implict               22.58  6  0.0009511 ***
SyntFun               59.10 12  3.287e-08 ***
Time                1503.53 42  < 2.2e-16 ***
CrChi.colloc.sig      94.26  6  < 2.2e-16 ***
CeChi.colloc.sig     206.99  6  < 2.2e-16 ***
CsedChi.colloc.sig   242.73  6  < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Prezi

	LR	Chisq	Df	Pr(>Chisq)	
Time	1503.53	42	<2.2e-16	***	
Manner	417.45	6	<2.2e-16	***	
CsedChi.colloc.sig	242.73	6	<2.2e-16	***	
CeChi.colloc.sig	206.99	6	<2.2e-16	***	
CsedSem	186.07	30	<2.2e-16	***	
CePers	103.79	36	1.72E-08	***	
CrChi.colloc.sig	94.26	6	<2.2e-16	***	
CrPers	70.32	36	0.000537	***	
CrSem	62.42	18	8.26E-07	***	
SyntFun	59.1	12	3.29E-08	***	
CeSem	39.49	18	0.002445	**	
CsedCstr	39.46	24	0.024406	*	
CsedNeg	25.05	6	0.000334	***	
Implict	22.58	6	0.000951	***	
Coref	20.15	12	0.064323	.	
CseNeg	14.11	6	0.028449	*	
CrDef	11.43	6	0.075949	.	
CeDef	10.57	6	0.102508		

Prezi

Levels: Significant?

```
rang vs shi

, , shi

                                2.5 %       97.5 %
(Intercept)              26.530350317  29.97251283
CrSemAnim                -0.840626523   0.18438545
...
CorefY                   -0.009365415   0.70745477
...
CsedCstrIntrans          -0.535536219   0.40467330
CsedCstrSVC              -1.125579358   1.07083698
CsedCstrTrans            -0.672091046   0.11261454
CsedSemEmotion           -0.573517155   0.69330695
...
```

✗ (In)direct causation hypothesis

STOP Ni (2012)

? · "shi"="doen", "rang"="laten"
· factors → different type of causality

Prezi

dCstrTrans	−0.672091046	0.1126
dSemEmotion	−0.573517155	0.6933

(In)direct causation hypothesis

Ni (2012)

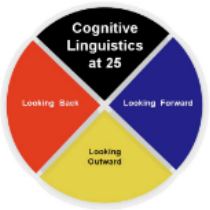
- **"shi"="doen", "rang"="laten"**
- **factors → different type of causality**

Prezi

What do we get already?

- There is CHANGE!
- It's hard to say which is stable.
  - occurrence
  - different complexions
- Witness of development
  - default form
  - (in)direct causation continuum

Prezi



for further information:
http://wwwling.arts.kuleuven.be/qlvl
yanan.hu@student.kuleuven.be
Dirk.Speelman@arts.kuleuven.be

Prezi