

What causes causatives?

A corpus-based typological study

Natalia Levshina

F.R.S.-FNRS, Université catholique de Louvain

Outline

1. Causative constructions
2. Data and variables
3. Quantitative analyses
4. Results and discussion

The causative continuum

Lexical <> Morphological <> Analytic (Periphrastic)

Comrie 1981, Dixon 2000, Shibatani & Pardeshi 2002, etc.

The causative continuum

Lexical <> Morphological <> Analytic (Periphrastic)

e.g. *kill, break*

The causative continuum

Lexical <> Morphological <> Analytic (Periphrastic)

e.g. *kill, break*

e.g. Turkish *öldür*- “kill”
from *öl*- “die”

e.g. *cause X to die,*
make X disappear

The causative continuum

Lexical <> Morphological <> Analytic (Periphrastic)

e.g. *kill, break*

e.g. Turkish *öldür*- “kill”
from *öl*- “die”

e.g. *cause X to die,*
make X disappear



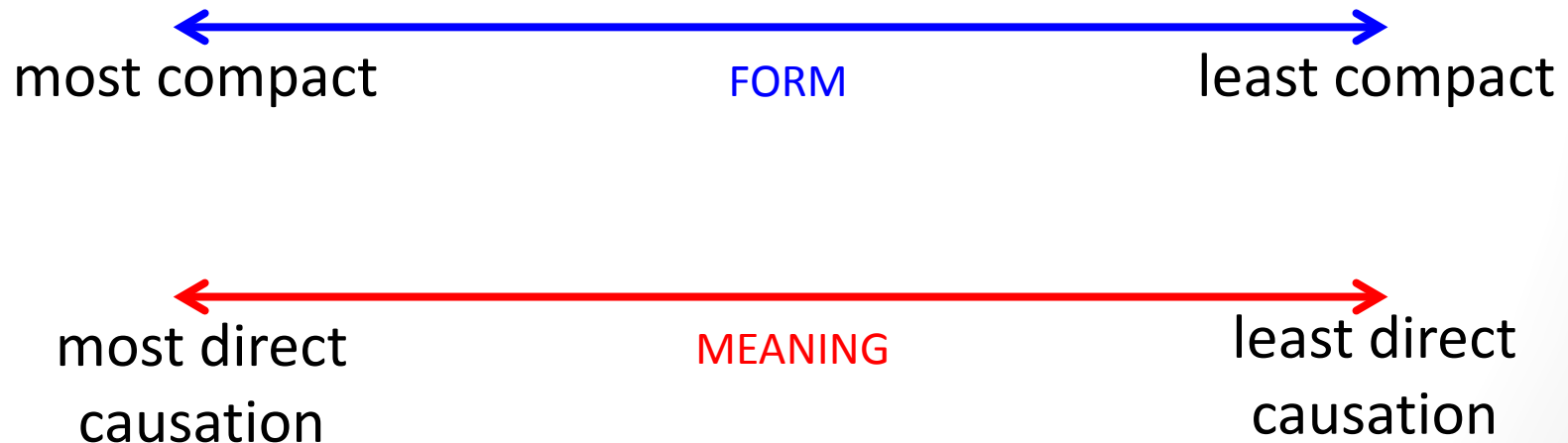
The causative continuum

Lexical <> Morphological <> Analytic (Periphrastic)

e.g. *kill, break*

e.g. Turkish *öldür*- “kill”
from *öl*- “die”

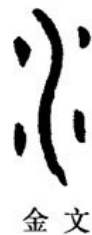
e.g. *cause X to die,*
make X disappear



Comrie 1981, Dixon 2000, Shibatani & Pardeshi 2002, etc.

Iconicity

- Most researchers agree that the degree of formal integration correlates with the degree of semantic integration of the cause and effect (e.g. Comrie 1981; Haiman 1983; Givón 1990) .
- An instance of iconic relationship between form and function.



Development of the Chinese character “water”

An extended approach

- Dixon (2000): a tentative list of 9 semantic and syntactic parameters based on a typological survey.
- Not all are directly interpretable in terms of iconicity.

Dixon's parameters



State (or change of state)
Intransitive
No control

**Relating to
VERB**

Action
(Di)transitive
Control

Willing ('let')
Partially affected

**Relating to
Causee**

Unwilling ('make')
Fully affected

Direct
Intentional
Natural

**Relating to
Causer**

Indirect
Accidental

With effort, violence

Dixon's parameters

More
compact

Less
compact



State (or change of state)
Intransitive

**Related to
base VERB**

Action
(Di)transitive

No control
Willing ('let')
Partially affected

**Related to
Causee**

Control
Unwilling ('make')
Fully affected

Direct
Intentional
Natural

**Related to
Causer**

Indirect
Accidental
With effort, violence

Dixon's parameters

More
compact

Less
compact



State (or change of state)
Intransitive

Related to
base VERB

Action
(Di)transitive

No control
Willing ('let')
Partially affected

Related to
Causee

Control
Unwilling ('make')
Fully affected

Direct
Intentional
Natural

Related to
Causer

Indirect
Accidental
With effort, violence

Dixon's parameters

More
compact

Less
compact



State (or change of state)
Intransitive
No control

Related to
VERB

Action
(Di)transitive
Control

Willing ('let')
Partially affected

Related to
Causee

Unwilling ('make')
Fully affected

Direct
Intentional
Natural

Related to
Causer

Indirect
Accidental
With effort, violence

The main question

- Can the formal variation (i.e. degree of compactness) of the causatives be explained by one factor (iconicity-related) or many factors (Dixon)?
- Never investigated quantitatively before!



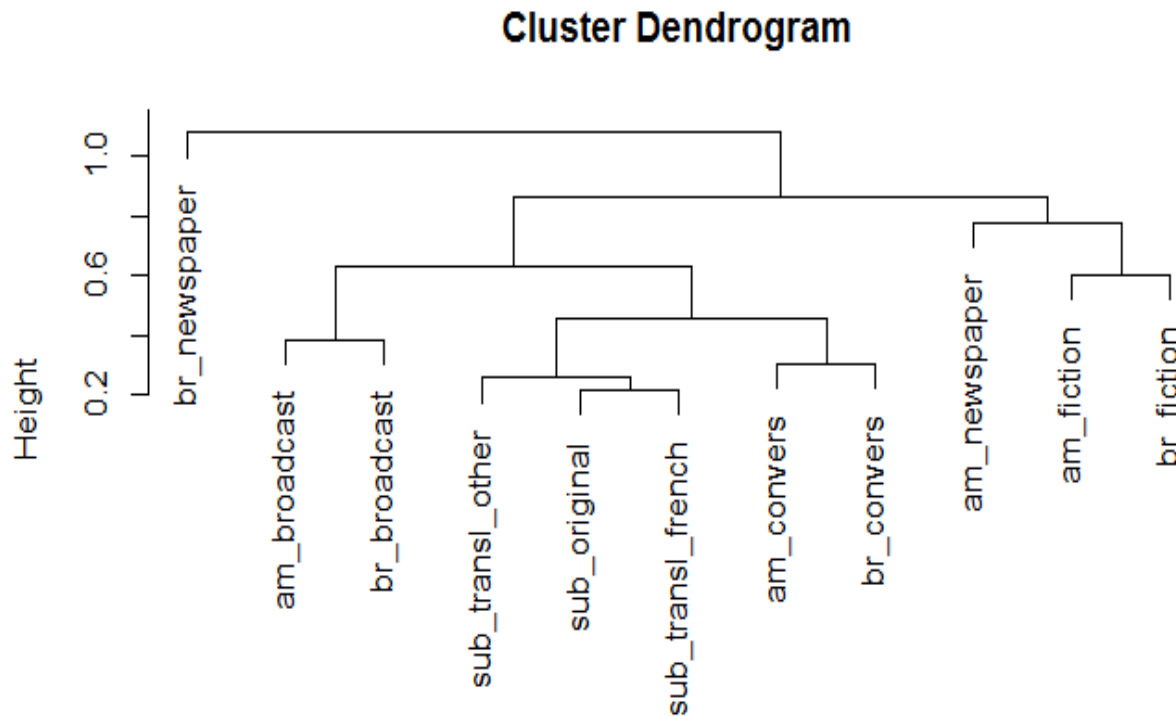
Outline

1. Causative constructions
2. **Data** and variables
3. Quantitative analyses
4. Results and discussion

ParTy corpus

- a Parallel corpus for Typology
- subtitles of films and TED talks
- mostly Indo-European languages, but also other major languages (Chinese, Turkish, Finnish, Indonesian, Japanese, Thai, etc.)
- all languages aligned with English
- downloadable files at www.natalialevshina.com/corpus.html
- work in progress...

Why subtitles?



Based on the frequencies of 3-grams (Levshina, Accepted)



Data used in the case study

Films



TED talks

- Ken Robinson: *Do schools kill creativity?*
- Elizabeth Gilbert: *Your elusive creative genius*
- Amy Cuddy: *Your body language shapes who you are*
- Leslie Morgan Steiner: *Why domestic violence victims don't leave*
- Dan Gilbert: *The psychology of your future self*
- Simon Sinek: *Why good leaders make you feel safe*

Languages

Language	Genus	Family
Chinese	Chinese	Sino-Tibetan
Finnish	Finnic	Uralic
French	Romance	Indo-European
Hebrew	Semitic	Afro-Asiatic
Indonesian	Malayo-Sumbawan	Austronesian
Japanese	Japanese	Japanese
Russian	Slavic	Indo-European
Thai	Kam-Tai	Tai-Kadai
Turkish	Turkic	Altaic
Vietnamese	Viet-Muong	Austro-Asiatic

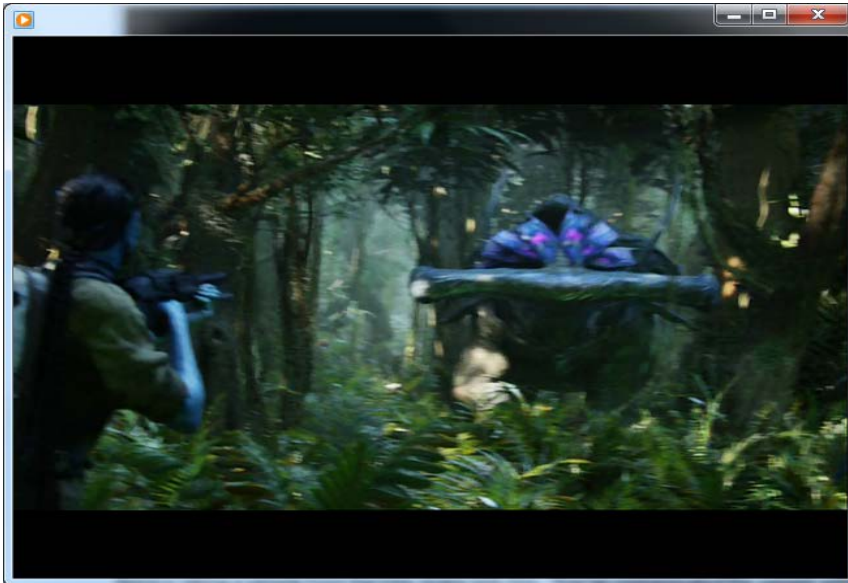
Data set

- 344 causative situations found in English
- Translations in the 10 languages are found and coded into 3 types of constructions (Analytic, Morphological or Lexical)

Example from *Avatar*

Original

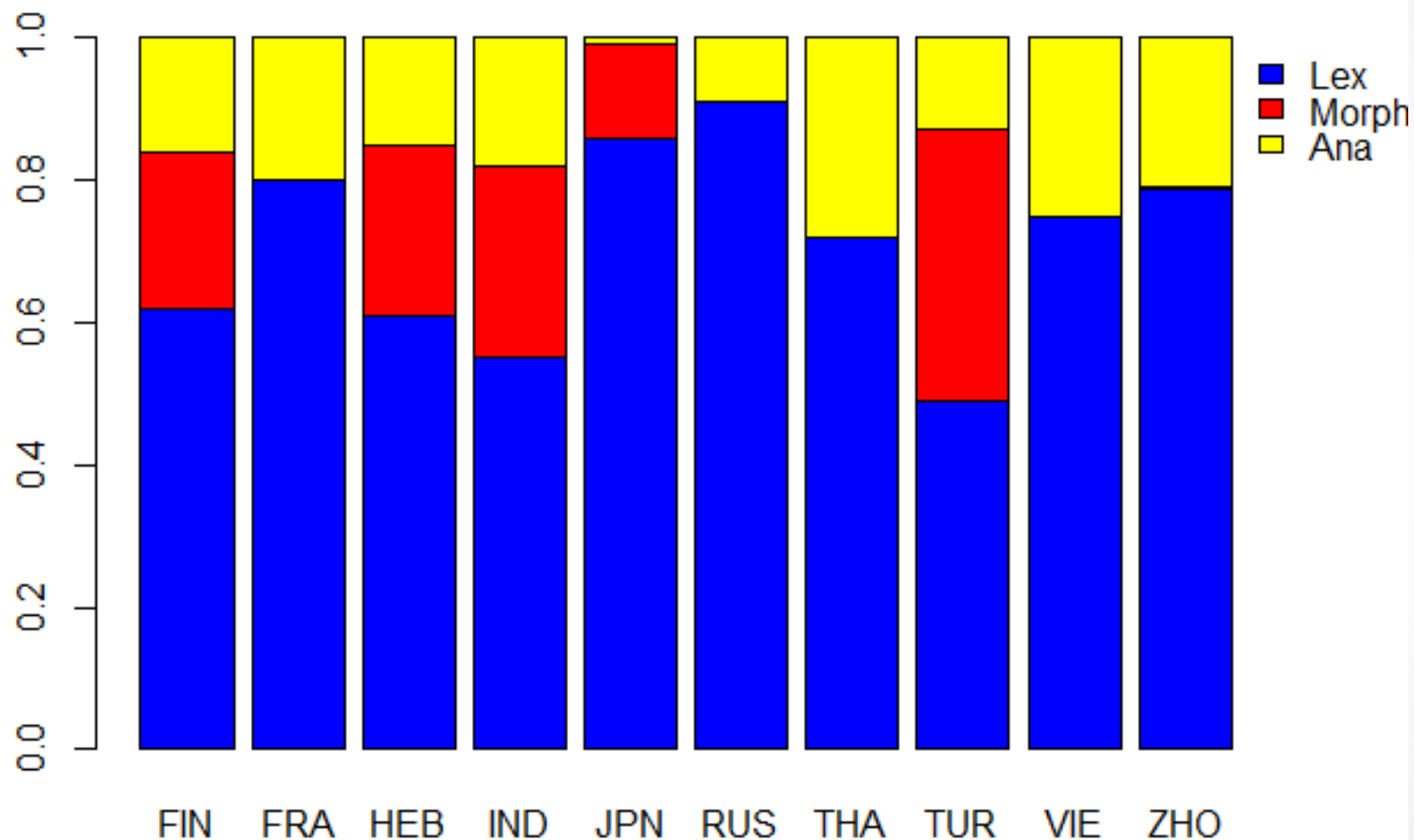
- ENG: *Don't shoot, you'll piss him off.*



Translations

- FRA: *Ne tirez pas. Vous allez l'énervé.* (Lexical)
- TUR: *Ateş etme. Ateş etme. Onu kızdıracaksın.* (Morphological, from *kızmek* 'become angry').
- VIE: *Đừng bắn. Cậu sẽ làm nó nổi điên đó.* (Analytic)

Proportions of types of causative constructions in 10 languages



Examples of constructions

	Lexical	Morphological	Analytic
Chinese	sha “kill”	-	ràng “let, make” + Pred
Finnish	tappaa “kill”	odotu-tt-aa “make wait”	antaa “give” + V1
French	tuer “kill”	-	faire + Vinf
Hebrew	harag “kill” <i>pa'al</i>	hotsi “take out” <i>hiph'il</i>	natan “give” + le-Vinf
Indonesian	mem-bunuh “kill”	meng-ingat-kan “remind”	membuat “make” + Pred
Japanese	korosu “kill”	ikar-ase-ru “make angry”	V_te + morau “get”
Russian	ubit' “kill”	-	zastavit' + Vinf
Thai	kaa “kill	-	tham hai “do give” + Pred
Turkish	açmak “open”	öl-dür- “kill”	V_mA_DAT + izin ver- “allow”
Vietnamese	giết hại “kill”	-	làm “do” + Pred

Data set

- 344 causative situations found in English
- Translations in the 10 languages are found and coded into 3 types of constructions (Analytic, Morphological or Lexical)
- The English sentences are coded for 13 semantic and syntactic variables (taking into account the context), based on Dixon's parameters...

Outline

1. Iconicity in causative constructions
2. Data and **variables**
3. Quantitative analyses
4. Results and discussion

Variables (1)

Variable	Values	Example(s)	Expectations
CausedEvent	Non-action Action	John killed Bill. He walked the baby upstairs.	Shorter form Longer form
NoPart (number of participants)	2 3	John killed Bill. I gave him a book.	Shorter form Longer form
CeControl (Causee having control)	No Yes	John killed Bill. Bring your friends!	Shorter form Longer form
MakeLet	Let Make	She let him go. John killed Bill.	Shorter form Longer form
CeVol (volitional Causee)	No Yes	John caused Bill to die. The police let him go.	Shorter form Longer form

Variables (2)

Variable	Values	Example(s)	Expectations
CrDirect (Causer acting directly)	Yes No	He cut his finger. He had his hair cut.	Shorter form Longer form
CrIntent (Causer acting intentionally)	Yes No	She wrote a paper. It makes me happy.	Shorter form Longer form
CrForce (Causer acting forcefully)	No Yes	He got him to do it. He forced him to do it.	Shorter form Longer form
CrInvolve (Causer involved in caused event)	No Yes	John killed Bill. Bring your friends! (and come, too)	None

Variables (3)

Variable	Values	Example(s)	Expectations
Coref (coreferentiality)	Yes No	He killed himself. He killed Bill.	None
Polarity	Pos Neg	She let him do it. She didn't let him do it.	None
CrSem (semantics of Causer)	Anim Inanim	She made him stay. The rain made him stay.	None
CeSem (semantics of Causee)	Anim Inanim	John let Mary go. John let it go.	None

Interrater agreement for semantic variables



Ludivine Crible, UCL



Samantha Laporte, UCL

Light's kappas

- Min = 0.7 *CrForce* (the Causer acting forcefully)
- Max = 0.93 *CrIntent* (the Causer acting intentionally)

Outline

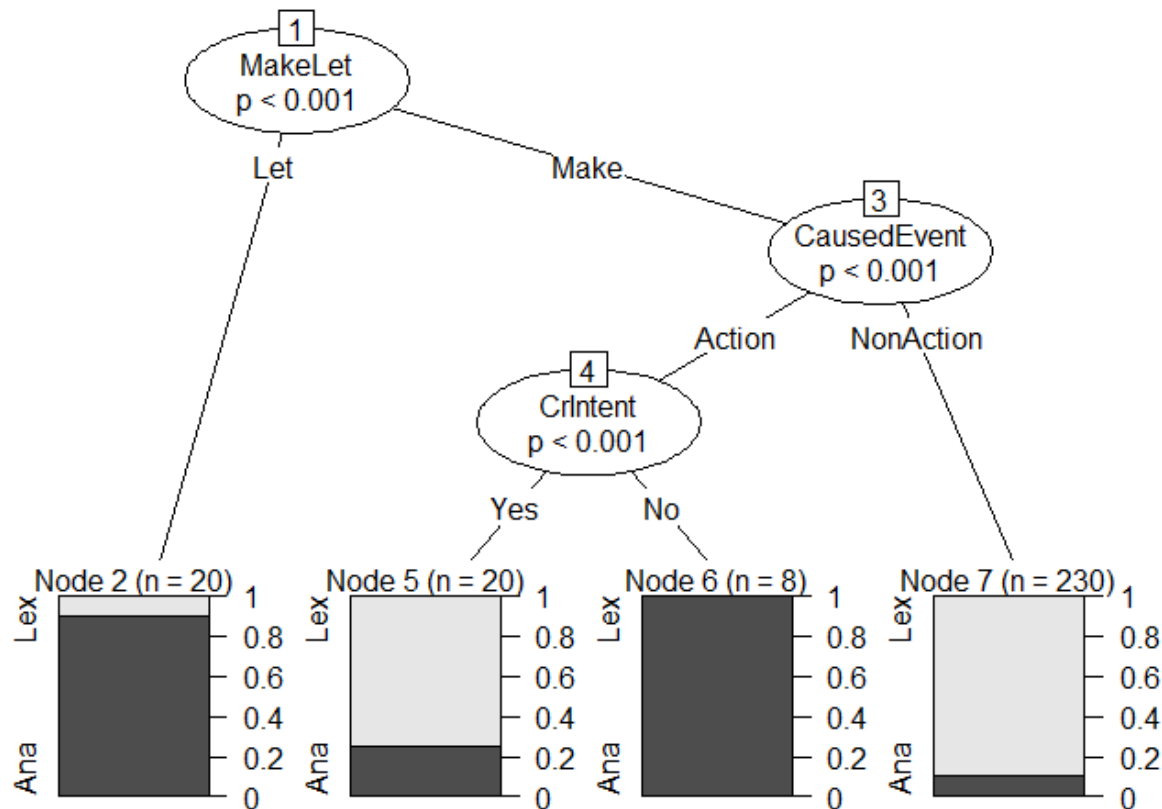
1. Causative constructions
2. Data and variables
3. **Quantitative analyses**
4. Results and discussion

Method: Random forests

- Statistical non-parametric method of classification (predicts the choice between causative constructions based on the 13 semantic and syntactic variables)
- Robust in situations of strongly correlated predictors and in situations of many predictors and few observations
- Based on many individual conditional trees



Example: a conditional tree for French (lexical vs. analytic)



Pourquoi ne pas les avoir **laissés** me **botter** le cul?

OK, **emmène-** moi avec toi.

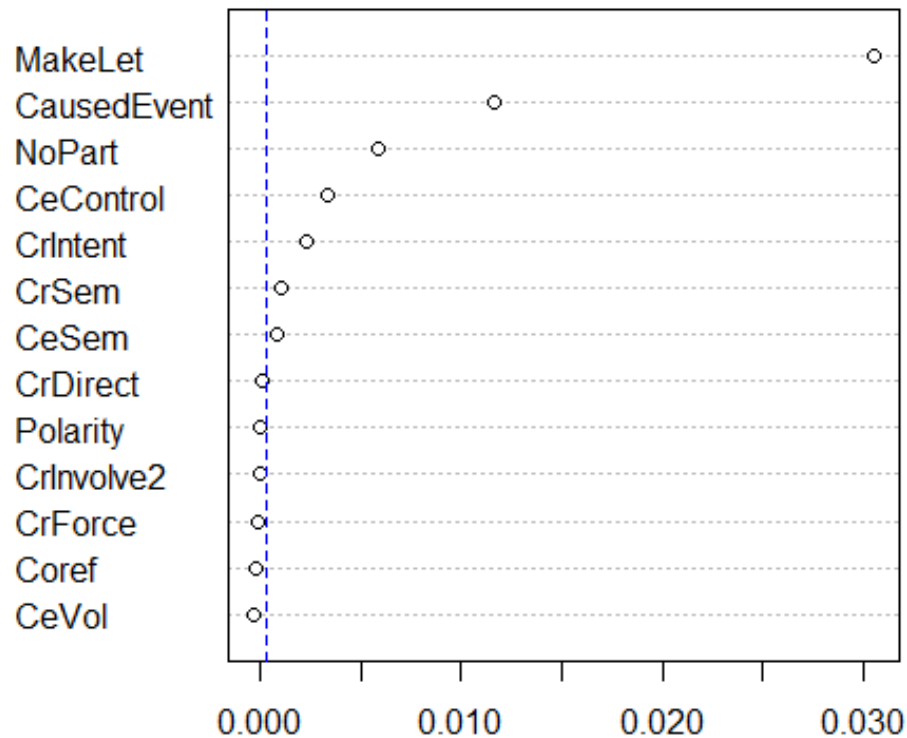
Tu sais ce que ça me **force à faire**?

Elle a **écrit** le livre.



Example: random forest for French (based on 1000 trees)

Variable importance



Var. importance: lex/ana

	FIN	FRA	HEB	IND	RUS	THA	TUR	VIE	ZHO
CausedEvent	High	High	High	High	High	High	High	High	High
NoPart	High	High	High	High	High	High	High	High	High
CeControl	High	High	High	High	Low	High	Low	High	High
MakeLet	Very High	Very High	Very High	Very High	Very High	Very High	Very High	Very High	Very High
CeVol	Low	Low	Low	High	Low	Low	Low	High	High
CrDirect	Low	Low	Low	Low	Low	Low	High	Low	Low
CrIntent	High	High	High	High	Low	High	Low	High	High
CrForce	Low	Low	Low	Low	Low	Low	High	Low	Low
CrInvolve	Low	Low	Low	High	Low	Low	Low	Low	Low
Coref	Low	Low	Low	Low	Low	Low	Low	Low	Low
Polarity	Low	Low	Low	Low	Low	Low	Low	Low	Low
CrSem	High	High	High	High	Low	High	Low	Low	High
CeSem	Low	Low	Low	High	Low	High	High	High	High

Var. importance: lex/morph

	FIN	HEB	IND	JPN	TUR
CausedEvent				High	
NoPart					
CeControl					
MakeLet				High	
CeVol		High			
CrDirect				High	
CrIntent	High		High	High	High
CrForce					
CrInvolve		High			
Coref					
Polarity					
CrSem	High			High	High
CeSem				High	High

Outline

1. Iconicity in causative constructions
2. Data and variables
3. Quantitative analyses
4. Results and discussion

Results

- Variation is clearly multivariate, not only cross-linguistically (Dixon), but also intra-linguistically. This may have to do with a variety of different causative constructions per language that have different functions.
- Variables related to semantic integration of cause and effect are the most prominent cross-linguistically, especially in the lex/ana variation:
 - Making or letting
 - CausedEvent: Action or non-action of the Causee
 - NoPart: Length of causation chain
 - Control of the Causee, etc.

Iconicity?

- Thus, in most cases we do observe iconic relationships: more compact forms express more semantically integrated events, less compact forms express less semantically integrated events.
- However, the Causer's (non)-intentionality is also important, especially in the lex/morph variation in Finnish, Indonesian and Turkish.

A usage-based explanation

- Principle of Economy(Haiman 1983; Haspelmath 2008).
- Indirect causation, as well as non-intentional causation, may be less frequent/familiar than direct and intentional causation = the core of the transitive prototype (Hopper & Thompson 1980).
- More complex forms are preferred to express less likely situations. This allows for efficient form-meaning mapping and equilibrium between speaker's and hearer's interests (Horn 1984). A result of language evolution based on cooperative social reasoning (Bergen, Goodman & Levi 2012).
- Is iconicity epiphenomenal?

Thank you!

The slides are available at

www.natalialevshina.com/presentations.html

Questions? Suggestions?

natalevs@gmail.com