Analytic causatives as comparative concepts: From Corpus-Driven to Traditional Semantic Maps

Natalia Levshina
F.R.S. – FNRS
Université catholique de Louvain
Outline

1. What are comparative concepts?
2. Analytic causatives
3. Data
4. Semantic maps
   - A probabilistic semantic map
   - A traditional semantic map
How to do typology?

• Typology tries to find generalizations of the type
  - “All languages have pronouns”
  - “If a language has the SOV word order, it will also have postpositions (e.g. I go the university to)“.

• To formulate such generalizations, one needs cross-linguistic categories, e.g. PRONOUN, POSTPOSITION, SUBJECT, OBJECT and VERB.
Problems

• The universal categories are problematic (there seems to be no such thing as a universal category ‘subject’ or ‘direct object’ wired in human genes).

• Language-specific descriptive categories are extremely diverse
  - English *go* is not like German *gehen* (go by bus?)
  - The German dative is not like the Russian dative (a present my friend-DAT?)

• But we need some basis for comparison (*tertium comparationis*)!
Comparative concepts (Haskelmath 2010)

- Technical categories created by linguists specially for purposes of language comparison.
- Do not have to correspond to descriptive categories in a specific language, like the German dative or Icelandic subject.
- Cannot be ‘right’ or ‘wrong’. Instead: more or less useful.
Outline

1. What are comparative concepts?
2. Analytic causatives
3. Data
4. Semantic maps
   - A probabilistic semantic map
   - A traditional semantic map
Analytic Causatives

MEANING

• An AC denotes a causative event, which consists of a causing event and a caused event.

• The main participants are the Causer, who initiates or is responsible for the causing event, and the Causee, who is the entity that brings about the caused event.

FORM

• An AC consists of two verbs and their arguments.

• V1 represents in an abstract way the causing event, and V2 represents the caused event.

• The order of V1 and V2 may vary.

• The Causee is the highest-ranking argument of V2; it is grammatically dependent on V1.
Causative events

*The magician made the rabbit disappear.*

**Causing event**

‘The **magician** does something’

**Causer**

**Caused event**

‘The **rabbit** disappears’

**Causee**
Analytic Causatives

**MEANING**
- AC denotes a causative event, which consists of a causing event and a caused event.
- The main participants are the causer, who initiates or is responsible for the causing event, and the causee, who is the entity that brings about the caused event (state).

**FORM**
- An AC consists of two verbs and their arguments.
- V1 represents in an abstract way the causing event, and V2 represents the caused event.
- The order of V1 and V2 may vary.
- The Causee is the highest-ranking argument of V2; it is grammatically dependent on V1.
Examples of ACs

• Eng. *make* + Vinf, *let* + Vinf
• Ger. *lassen* + Vinf, *bringen zu* + V
• Rus. *zastavljat’* ‘force’ + Vinf, *pozvoljat’* ‘allow’ + Vinf
• Fin. *antaa* ‘give, let’ + 1\textsuperscript{st} Infinitive
• Rom. *a face* ‘make’ + sǎ VSubj
• Bul. *karam* ‘drive, make’ + *da* Vfin
Excluded (examples)

- Eng. *into*-causative: *The neighbour fooled her into thinking that her camera causes accidents*. Too specific V1.
- Rus. *sdelat’ tak, čtoby* ... ‘do it so that’. Causee is not dependent on V1.

\[Ja\ \text{sdelal\ tak,\ čtoby\ on\ byl\ postojanno\ rjadom.}\]
\[I\ \text{did\ so\ COMP\ he\ was\ always\ near}\]
\[‘I\ made\ it\ so\ that\ he\ was\ always\ near.’\]
Languages
Outline

1. What are comparative concepts?
2. Analytic causatives
3. Data
4. Semantic maps
   - A probabilistic semantic map
   - A traditional semantic map
Parallel corpus of film subtitles
An example of .srt format

... 646 00:51:27,880 --> 00:51:32,920
<i>For always evil will look to find a foothold in this world.</i></p> 647 00:51:39,440 --> 00:51:42,603
Not good. Not good at all.

648 00:51:50,040 --> 00:51:51,326
Eww.

649 00:52:06,760 --> 00:52:09,081
Oh, no. Sebastian.

650 00:52:12,800 --> 00:52:13,847
Good gracious.

651 00:52:34,720 --> 00:52:35,767
Come on.

...
Data set

- 1251 instances of ACs in 18 languages
- 390 multilingual contexts (with equivalent sentences in 18 languages), where at least one language contains an AC
Frequencies of ACs
Coding of ACs

- V1 (*make, faire, lassen, davat’, etc.*)
- Form of V2 (infinitive, finite verb, past participle, etc.)
- Intransitive or transitive V2 (e.g. *The magician made the rabbit disappear* vs. *The magician let the rabbit eat a carrot.*)
- Coreferential or non-coreferential with Causer (e.g. *Lass dich nicht erwischen!* vs. *Lass ihn gehen.*)
- Passive or active V1 (e.g. *They forced him to do it* vs. *He was forced to do it.*)
Outline

1. What are comparative concepts?
2. Analytic causatives
3. Data
4. Semantic maps
   - A probabilistic semantic map
   - A traditional semantic map
Semantic maps

- a popular tool in constructional (lexical) and semantic typology
- represent polyfunctionality of linguistic expressions
- a convenient tertium comparationis for form-function mapping across languages
- Come in two main flavours: traditional (semantic functions and links) and probabilistic (exemplars and distances)
Traditional semantic map of datives (Haspelmath 2003)

The English Dative preposition *to*
the French **Dative preposition** à

- predicative possessor
- external possessor
- beneficiary
- judicantis
- direction
- recipient
- experiencer
- purpose

Haspelmath 2003: Datives
Main principles

- Nodes: A function is put on a map when there’s at least one pair of languages which differ wrt. this function (Haspelmath 2003)
- Links: the principle of connectivity (adjacency/contiguity):
  if a construction has more than one function, they should be connected (see van der Auwera 2013)
Example 1
Example 2

Function A

Function B

Function C

Construction Y
Example 3

Function A

Function B

Function C

Construction Z
Example 4

Wrong: the connectivity principle is not observed!
A fix
Probabilistic maps

• Display distances and exemplars
• Usually based on parallel corpus data
• Statistical techniques (Multidimensional Scaling)
• Main advantage: help identify clusters (semantic functions) and dimensions of semantic variation. This comes in handy when we do not know in advance what the functions are like.
Wälchli & Cysouw (2012): motion verbs
Outline

1. What are comparative concepts?
2. Analytic causatives
3. Data
4. Semantic maps
   - A probabilistic semantic map
   - A traditional semantic map
Situation (row) A
EN: And we make them do it... ...or we kill them.
   make_Vinf_Trans_NoCoref_Act
IT: E glielo facciamo fare ... o lo uccidiamo.
   fare_Vinf_Trans_NoCoref_Act
RO: Apoi îl obligăm să o facă ... sau îl ucidem ..
   obliga săVsubj_Trans_NoCoref_Act

Situation (row) B
EN: Pick up someone my height and build and make them believe it is me.
   make_Vinf_ComplClause_NoCoref_Act
IT: Individua una della mia corporatura e fa credere loro che sia io.
   fare_Vinf_Trans_NoCoref_Act
RO: Alege pe cineva care seamănă cu mine și fă- i să creadă că sunt eu.
   face săVsubj_Trans_NoCoref_Act
Multidimensional Scaling

- Multidimensional Scaling of the distance matrix.
- The main principle: the closer two points on the map, the more overlapping constructions they share across the languages. From the isomorphism principle it follows that the corresponding situations are more semantically similar (on average), since more authors of the subtitles chose identical constructions to represent these causative events.
Letting vs. making

**LETTING**
- More systematically covered by ACs across the languages
- Less formal variation
- Simpler V2 forms

**MAKING**
- Less systematically covered by ACs across the languages
- More formal variation
- More complex V2 forms
MAKING: Competition with lexical causatives

French: analytic
Amandine Poulain aime: (...) Faire briller le parquet...
Amandine Poulain likes: (...) make shine the parquet
‘Amandine Poulain likes: (...) polishing the parquet’

English: lexical (transitive verb)
Amandine Poulain likes: (...) polishing the parquet...
MAKING: Competition with morphological causatives

• French: analytic

_Le vent s'engouffrait comme par magie sous une nappe_,
the wind REFL-plunged like by magic under a tablecloth,
*faisant danser les verres*...
making dance the glasses
‘The wind crept under a tablecloth, as if by magic, making the glasses dance.’

• Finnish: morphological

..._tuuli tanssi-tt-i kahta las-ia pöytäliin-alla_*
wind dance-CAUS-PST.3SG two glass-PART.SG tablecloth-AD.SG
‘...the wind made two glasses dance on the tablecloth.’
Germanic ‘sandwich’

- English > Dutch > German: a grammaticalization cline. English is further down on the cline than Dutch and especially German in many areas of grammar, e.g. loss of gender, case, verb endings (e.g. Hüning, Vogel, van der Wouden & Verhagen 2006).
LET-ACs in Germanic
Romance ‘sandwich’

- De Mulder & Lamiroy 2012: Romance languages exhibit different degrees of grammaticalization in different areas of grammar, as well:

  French > Italian > Spanish
MAKE-ACs in Romance
Italian *fare* + Vinf: letting

*Hanno fermato il treno per far-la salire.*

have stopped the train for make-her mount

‘They stopped the train to let her get on.’

*Stronzate, fammi vedere!*

bullshit make.me see

‘Bullshit, let me see that!’
GIVE-causatives in Slavic

• Letting, e.g. Russian:

\[ Daj-te \text{ mne, požalujsta, pospat’}. \]

\text{give.IMP.PFV-2PL me please sleep.INF.PFV}

‘Please, let me sleep!’

• Curative, e.g. Slovenian:

\[ Dal \text{ ga bom nadzorovati.} \]

\text{give.PASTPART him be.FUT.1SG watch.IPFV.INF}

‘I’ll have him monitored.’

• Middle voice, e.g. Czech:

\[ Uvidíme, jestli se tvé bláznovství dá vyléčit. \]

\text{see.PFV.1PL if REFL your insanity give.PFV.3SG cure.INF.PFV}

‘And we will see if your insanity can be cured.’
Slavic ‘soup’

ACs with GIVE auxiliaries in Slavic
Language contact as a major factor of variation of ACs

• Czech and Slovenian: GIVE-causatives express curative causation
  influence of *lassen* + Vinf (cf. von Waldenfels 2012)
• Estonian: *laskma* + Vda – very similar to *lassen* + Vinf
  again, it seems to be influence of German (or Swedish)
• Semantic bleaching of letting and making ACs in the core of Standard Average European (Charlemagne Sprachbund)
  are highly frequent, strongly grammaticalized and semantically bleached ACs an areal property?
Charlemagne Sprachbund

- van der Auwera 1998
- the area of contemporary Germany, the Netherlands, Belgium, France and northern Italy
- Intensive contact since appr. the 9th century
- The core of Standard Average European

Karl der Große (? - 814), the founder of the Carolingian Empire
Growth of Frankish power (Shepherd 1926)
Outline

1. What are comparative concepts?
2. Analytic causatives
3. Data
4. Semantic maps
   - A probabilistic semantic map
   - A traditional semantic map
A pilot study

- 7 languages: English, German, French, Portuguese, Finnish, Czech, Russian
- Matrix with 6 functions and 44 ACs from the corpus (Present = 1, Absent = 0)
- R code that generates possible semantic maps by reshuffling the functions
- A neglected issue in typology: Many possible configurations
Thanks!

Rus. ‘We’ll make you smile!’