Analytic causatives in European languages:

A quantitative study based on a multilingual parallel corpus

Natalia Levshina
F.R.S. – FNRS
Université catholique de Louvain
Aims of this study

• create a probabilistic semantic map of analytic causatives (ACs) in 12 European languages
• compare the form-function mappings of language-specific ACs with the help of the map and identify genealogical and geographic patterns
Outline

1. Analytic causatives as a comparative concept
2. Data and method
3. Probabilistic semantic space of European ACs
4. Geographic and genealogical variation in form-meaning mapping
5. Conclusions
Linguistic categories and language comparison

- Language-specific descriptive categories (structuralism and CxGr)
  → impossible to compare directly!
Linguistic categories and language comparison

- Language-specific descriptive categories (structuralism and CxGr)
  → impossible to compare directly!

- Universal cross-linguistic categories (generative typology)
  → theoretically problematic (Croft 2001) and practically unsuccessful (e.g. Evans & Levinson 2009) (at least, so far...)

Linguistic categories and language comparison

• Language-specific descriptive categories (structuralism and CxGr)
  → impossible to compare directly!

• Universal cross-linguistic categories (generative typology)
  → theoretically problematic (Croft 2001) and practically unsuccessful (e.g. Evans & Levinson 2009) (at least, so far...)

• Solution: comparative concepts (cf. Haspelmath 2010)...
Comparative concepts

• “concepts created by comparative linguists for the specific purpose of language comparison. Unlike descriptive categories, they are not part of particular language systems and are not needed by descriptive linguists or by speakers. They are not psychologically real, and they cannot be right or wrong. They can only be more or less well-suited to the task of permitting cross-linguistic comparison” (Haspelmath 2010)
Comparative concepts

• “concepts created by comparative linguists for the specific purpose of language comparison. Unlike descriptive categories, they are not part of particular language systems and are not needed by descriptive linguists or by speakers. They are not psychologically real, and they cannot be right or wrong. They can only be more or less well-suited to the task of permitting cross-linguistic comparison” (Haspelmath 2010)
• Contain form and function (except phonology) and tend to include other comparative concepts, e.g. WORD, VERB or CLAUSE (Haspelmath 2010).
Comparative concepts

- “concepts created by comparative linguists for the specific purpose of language comparison. Unlike descriptive categories, they are not part of particular language systems and are not needed by descriptive linguists or by speakers. They are not psychologically real, and they cannot be right or wrong. They can only be more or less well-suited to the task of permitting cross-linguistic comparison” (Haspelmath 2010).

- Comparative concepts are constructions, which may but do not have to correspond to constructions in any specific language.

- Comparative concepts contain form and function (except phonology) and tend to include other comparative concepts, e.g. WORD, VERB or CLAUSE (Haspelmath 2010).
Analytic causatives as Comparative Concepts

- An AC is a construction that...
Analytic causatives as Comparative Concepts

• An AC is a construction that...
• designates a causative event, which involves a causing and caused events
Analytic causatives as Comparative Concepts

• An AC is a construction that...
• designates a causative event, which involves a causing and caused events
• The obligatory core of an AC consists of two VERBS:
Analytic causatives as Comparative Concepts

• An AC is a construction that...
• designates a causative event, which involves a causing and caused events
• The obligatory core of an AC consists of two VERBS:
  • one VERB (V1) represents in an abstract way the causing event
Analytic causatives as Comparative Concepts

• An AC is a construction that...
• designates a causative event, which involves a causing and caused events
• The obligatory core of an AC consists of two VERBS:
  • one VERB (V1) represents in an abstract way the causing event
  • the other VERB (V2) represents the caused event
Analytic causatives as Comparative Concepts

- An AC is a construction that...
- designates a causative event, which involves a causing and caused events
- The obligatory core of an AC consists of two VERBS:
  - one VERB (V1) represents in an abstract way the causing event
  - the other VERB (V2) represents the caused event
  - V2 is grammatically dependent on V1
Analytic causatives as Comparative Concepts

• An AC is a construction that...
• designates a causative event, which involves a causing and caused events
• The obligatory core of an AC consists of two VERBS:
  • one VERB (V1) represents in an abstract way the causing event
  • the other VERB (V2) represents the caused event
  • V2 is grammatically dependent on V1
  • The semantic subject of V2 (the Causee) is an OBJECT of V1, expressed in the surface structure or implied.
Illustrations

- EN: *make* + V

  *Don’t make me cry, babe.*
Illustrations

• EN: make + V

  Don’t make me cry, babe.
Illustrations

• EN: make + V
  
  Don’t make me cry, babe.

• EN: order + to-V
  
  The general ordered the troops to shoot.
Illustrations

• EN: *make* + V

  Don’t make me cry, babe.

• EN: *order* + to-V

  The general ordered the troops to shoot.
Illustrations

• EN: make + V
  
  *Don’t make me cry, babe.*

• EN: order + to-V
  
  *The general ordered the troops to shoot.*

• Atchin: “AND-type” (Song 1996)

  mar kete ni-wat mu tsöv

  3PL.PST make stone 3SG.PST fall
  
  “They made the stone fall“.
Illustrations

• EN: *make* + V
  
  *Don’t make me cry, babe.*

• EN: *order* + *to-V*
  
  *The general ordered the troops to shoot.*

• Atchin: “AND-type” (Song 1996)
  
  *mar kete ni-wat mu tsöv*
  
  3PL.PST make stone 3SG.PST fall
  
  “They made the stone fall“.
Illustrations

• EN: make + V

  Don’t make me cry, babe.

• EN: order + to-V

  The general ordered the troops to shoot.

• Atchin: “AND-type” (Song 1996)

  mar kete ni-wat mu tsöv
  3PL.PST make stone 3SG.PST fall
  “They made the stone fall“.

• RU: (s)delat‘ tak, čtoby ... “make so that”

  Ja sdelal tak, čtoby on byl postojanno rjadom.
  “I made it so that he was always near.”
Illustrations

• EN: make + V
  
  *Don’t make me cry, babe.*

• EN: order + to-V
  
  *The general ordered the troops to shoot.*

• Atchin: “AND-type” (Song 1996)
  
  *mar kete ni-wat mu tsöv*
  
  3PL.PST make stone 3SG.PST fall
  
  “They made the stone fall“.

• RU: (s)delat‘ tak, čtoby … “make so that”
  
  *Ja sdelal tak, čtoby on byl postojanno rjadom.*
  
  “I made it so that he was always near.”
# Methods in constructional typology

<table>
<thead>
<tr>
<th>Method</th>
<th>Example</th>
<th>Context</th>
<th>Linguist-independent ‘senses’</th>
</tr>
</thead>
<tbody>
<tr>
<td>Decompositional method</td>
<td>Wierzbicka’s NSM</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Questionnaires and wordlists</td>
<td>CLICS database (List et al. 2014)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Experiments (only perceptual!)</td>
<td>MPI Nijmegen</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Analysis of parallel corpora</td>
<td>Wälchli &amp; Cysouw 2012</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>
### Methods in constructional typology

<table>
<thead>
<tr>
<th>Method</th>
<th>Example</th>
<th>Context</th>
<th>Linguist-independent ‘senses’</th>
</tr>
</thead>
<tbody>
<tr>
<td>Decompositional method</td>
<td>Wierzbicka’s NSM</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Questionnaires and wordlists</td>
<td>CLICS database (List et al. 2014)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Experiments (only perceptual!)</td>
<td>MPI Nijmegen</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Analysis of parallel corpora</td>
<td>Wälchli &amp; Cysouw 2012</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>
DATA & QUANTITATIVE METHODS
Multilingual 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>

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

- All instances of ACs in 12 languages
- 362 multilingual contexts, in which at least one doculect contains an AC
- Alignment: Jörg Tiedemann’s software subalign
Constructional types

- ACs with various auxiliaries (e.g. *lassen* + V, *fazer* + V)
- Transitives and ditransitives
- Causative verb + Clause (e.g. *dejar* + *que*)
- Causal prepositions (*because of*)
- Causal and resultative subordinate clauses
- Resultative cxs, e.g. *make* + Adj
- Modals
- Particles (*niech* in Polish)
- Reflexives
- Insubordination (*que* + Subj)
- ...
### Matrix

<table>
<thead>
<tr>
<th>FR</th>
<th>EN</th>
<th>DE</th>
<th>ES</th>
<th>NL</th>
<th>SV</th>
<th>IT</th>
</tr>
</thead>
<tbody>
<tr>
<td>faire_V</td>
<td>Trans</td>
<td>Trans</td>
<td>Trans</td>
<td>Trans</td>
<td>Trans</td>
<td>fare_V</td>
</tr>
<tr>
<td>faire_V</td>
<td>Trans</td>
<td>NA</td>
<td>Trans</td>
<td>Trans</td>
<td>Trans</td>
<td>fare_V</td>
</tr>
<tr>
<td>faire_V</td>
<td>have_Ved</td>
<td>lassen_V</td>
<td>Trans</td>
<td>Trans_Pass</td>
<td>Trans</td>
<td>fare_V</td>
</tr>
<tr>
<td>faire_V</td>
<td>lassen_V</td>
<td>lassen_V</td>
<td>dejar_V</td>
<td>laten_V</td>
<td>lata_V</td>
<td>fare_V</td>
</tr>
<tr>
<td>faire_V</td>
<td>lassen_V</td>
<td>lassen_V</td>
<td>CausClause</td>
<td>ResClause</td>
<td>ResClause</td>
<td>fare_V</td>
</tr>
<tr>
<td>faire_V</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>NA</td>
<td>fare_V</td>
</tr>
<tr>
<td>faire_V</td>
<td>Trans_Pass</td>
<td>NA</td>
<td>Trans</td>
<td>Trans_Pass</td>
<td>Trans_Pass</td>
<td>fare_V</td>
</tr>
<tr>
<td>faire_V</td>
<td>lassen_V</td>
<td>lassen_V</td>
<td>doen_V</td>
<td>fa_att_V</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>faire_V</td>
<td>have_Ved</td>
<td>lassen_V</td>
<td>NA</td>
<td>laten_V</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>faire_V</td>
<td>lassen_V</td>
<td>lassen_V</td>
<td>TransPass</td>
<td>NA</td>
<td>Trans</td>
<td>fare_V</td>
</tr>
<tr>
<td>faire_V</td>
<td>NA</td>
<td>NA</td>
<td>Trans</td>
<td>NA</td>
<td>Trans</td>
<td>fare_V</td>
</tr>
<tr>
<td>faire_V</td>
<td>NA</td>
<td>NA</td>
<td>Trans</td>
<td>NA</td>
<td>Trans</td>
<td>fare_V</td>
</tr>
<tr>
<td>faire_V</td>
<td>Trans_Pass</td>
<td>Trans_Pass</td>
<td>NA</td>
<td>Trans_Pass</td>
<td>NA</td>
<td>fare_V</td>
</tr>
<tr>
<td>faire_V</td>
<td>lassen_V</td>
<td>lassen_V</td>
<td>deja_r</td>
<td>Intrans</td>
<td>Intrans</td>
<td>fare_V</td>
</tr>
<tr>
<td>faire_V</td>
<td>lassen_V</td>
<td>lassen_V</td>
<td>TransPass</td>
<td>Trans_Pass</td>
<td>Refl</td>
<td>fare_V</td>
</tr>
<tr>
<td>faire_V</td>
<td>lassen_V</td>
<td>lassen_V</td>
<td>dejar_V</td>
<td>laten_V</td>
<td>lata_V</td>
<td>lasciare_V</td>
</tr>
<tr>
<td>faire_V</td>
<td>lassen_V</td>
<td>lassen_V</td>
<td>dejar_V</td>
<td>laten_V</td>
<td>lata_V</td>
<td>lasciare_V</td>
</tr>
<tr>
<td>Trans</td>
<td>Trans</td>
<td>lassen_V</td>
<td>Trans</td>
<td>laten_V</td>
<td>NA</td>
<td>Trans</td>
</tr>
<tr>
<td>Trans</td>
<td>Trans</td>
<td>Trans</td>
<td>laten_V</td>
<td>NA</td>
<td>Trans</td>
<td>Trans</td>
</tr>
</tbody>
</table>
Gower’s Distances

Situation (row) A
EN: And we make them do it... ...or we kill them. make_V
IT: E glielo facciamo fare ... o lo uccidiamo. fare_V
CZ: Donutíme je to udělat, nebo je zabijeme. donutil_V

Situation (row) B
EN: Pick up someone my height and build and make them believe it is me. make_V
IT: Individua una della mia corporatura e fa credere loro che sia io. fare_V
CZ: Vyber někoho, kdo je mi podobný a přesvědč je, že jsem to já. Trans

Distance (A, B) = 1 – 2/3 ≈ 0.33
Statistical methods

- Multidimensional Scaling of the distance matrix.
Statistical methods

• 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 doculects chose identical constructions to represent these causative situations.
Statistical methods

• 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 doculects chose identical constructions to represent these causative situations.
• Therefore, the result is a probabilistic semantic map (Wälchli & Cysouw 2012).
PROBABILISTIC SEMANTIC MAP OF ACs
MDS-based semantic map

Probabilistic semantic map of ACs
FR: Alors après avoir tué ce clébard à la botte des allemands... on ira les faire taire une bonne fois pour toutes.

EN: Then after we kill this dog for the Germans ... we'll go and silence them.
Letting (non-interference)

EN: I honestly didn't think he would let it go this far.
Letting (removal of blockage)

EN: They let me go.
Indirect causation (distance or agentive Ce)

EN: I can't believe you're making me do this.
Passive (Causer = Affectee)

FR: Ne te fais pas avoir.
EN: Don’t you fall for that.
EN: To you! To letting all of Germany see the true face of the GDR!
DE: Auf dass Sie Gesamtdeutschland das wahre Gesicht der DDR zeigen.
CROSS-LINGUISTIC VARIATION IN FORM-MEANING MAPPING
All ACs: Romance & English

[French ACs, Italian ACs, Spanish ACs, Portuguese ACs, Romanian ACs, English ACs]
All ACs: Germanic & Slavic
ACs with ‘make’: Romance & English
ACs with ‘let’: Romance & English
ACs with ‘let’: Germanic & Slavic
CONCLUSIONS
Conclusions

• Unlike Direct Causation and Letting (Non-impingement), Indirect Causation does not form a dense and distinct cluster. This means that the languages exhibit great variation in verbalization of this semantics in comparison with the other major senses.
Conclusions

• Unlike Direct Causation and Letting (Non-impingement), indirect Causation does not form a dense and distinct cluster. This means that the languages exhibit great variation in verbalization of this semantics in comparison with the other major senses.

• The mappings suggests the following coarse-grained hierarchy of semantic functions of ACs as abstract types in this set of languages:

  letting < indirect causation < direct causation
Conclusions

• Unlike Direct Causation and Letting (Non-impingement), Indirect Causation does not form a dense and distinct cluster. This means that the languages exhibit great variation in verbalization of this semantics in comparison with the other major senses.

• The mappings suggests the following coarse-grained hierarchy of semantic functions of ACs as abstract types in this set of languages:

  letting < indirect causation < direct causation

• Although the plots reveal some genealogical similarities, this does not hold for all languages and language groups. In addition, some patterns can be best explained by language contact (e.g. CZ and DE).
A hypothesis

- German *lassen* + V, Dutch *laten* + V, French *faire* + V and Italian *fare* + V are the most prominent quantitatively and the most diverse (‘empty’) qualitatively.
A hypothesis

- German *lassen* + V, Dutch *laten* + V, French *faire* + V and Italian *fare* + V are the most prominent quantitatively and the most diverse (‘empty’) qualitatively.
- Since these languages constitute the nucleus of Standard Average European (Charlemagne Sprachbund), are such ACs a SAE feature? (cf. van der Auwera 1998; Haspelmath 2001)
He is not a big fan of causatives...
In Slovene: variation between Vinf after V1 and 
*da*-clause:
- *dati* “give” + Vinf: *Dal mi je vedeti, da sem slab študent* “He let me know that I am a bad student”. (Greenberg 2006: 131) AC
- *dati* + *da* + V<sub>Indicative</sub>: *Zajebavaš, daj da vidim.* “You’re kidding, let me see.” NOT AC
(to be explored)
Non-finiteness = ‘desententialization’, i.e. the degree of downgrading of the dependent clause with respect to the prototypical independent/main clause.

- gradual and language-specific
- Relatively uncontroversial features: no tense marking, lack of subject agreement (Nikolaeva 2007)
MCA of abstract AC types
MCA of specific types
Analytic causatives as Comparative Concepts

• An AC is a construction that...
• designates a causative event, which involves a causing and caused events
• The obligatory core of an AC consists of two VERBS:
  • one VERB (V1) represents in an abstract way the causing event
  • the other VERB (V2) represents the caused event
  • V2 is grammatically dependent on V1
  • The semantic subject of V2 (the Causee) is an OBJECT of V1, expressed in the surface structure or implied.

VERB: A WORD that (typically) expresses predication (vs. reference and modification) and action (vs. objects and states) (Croft 1991).
Analytic causatives as Comparative Concepts

- An AC is a construction that...
- designates a causative event, which involves a causing and caused events
- The obligatory core of an AC consists of two VERBS:
  - one VERB (V1) represents in an abstract way the causing event
  - the other VERB (V2) represents the caused event
- V2 is grammatically dependent on V1
- V2 is grammatically expressed in the surface structure or implied.

VERB: A WORD that (typically) expresses predication (vs. reference and modification) and action (vs. objects and states) (Croft 1991).

WORD: “a segment string that cannot be interrupted by a free form without changing the meaning” (Haspelmath 2010: 4)