

Dative Subjects: Historical Change Visualized

Christin Schätzle

Workshop on Word Order in Scandinavian languages
University of Konstanz

December 14th, 2018

- ▶ Investigation of the interaction between **dative subjects**,
 - ▶ lexical semantics,
 - ▶ event structure,
 - ▶ voice (active, passive, middle),
 - ▶ and word order

in the Icelandic Parsed **Historical** Corpus (IcePaHC).

- ▶ Aim: Shed light on the function of case marking in the complex system which licenses grammatical relations in the history of Icelandic.
- ▶ Method of investigation: **Visual Analytics** for historical linguistics (HistLingVis)
- ▶ Contribution to the on-going discussion on whether dative subjects are a Proto Indo-European inheritance

- ▶ 12th to 21st century – all attested stages of Icelandic.
- ▶ 61 texts, 1 million words, different genres (not representative across centuries).
- ▶ Approximately 100 000 words per century.
- ▶ Annotation based on Penn Treebank-style (Marcus et al. 1993).
- ▶ Information about sentence types, constituents, word order, grammatical relations, tense, voice, and case.

Sample IcePaHC Annotation

```
(IP-MAT-SPE (NP-SBJ (PRO-D Mér-mér))
  (VBPI finnst-finna)
  (CP-ADV-SPE (WADV-1 0)
    (C sem-sem)
    (IP-SUB-SPE (ADVP *T*-1)
      (NP-SBJ (PRO-N ég-ég))
      (BEPS sé-vera) (VBN sloppinn-sleppa)
      (PP (P úr-úr) (NP (NP-POS (ONE+Q-G einhvers-einhver)
        (N-G konar-konar)) (N-D fangelsi-fangelsi))))))
  (. .-.))
(ID 1882.TORFHILDUR.NAR-FIC,.603))
```

Data processing

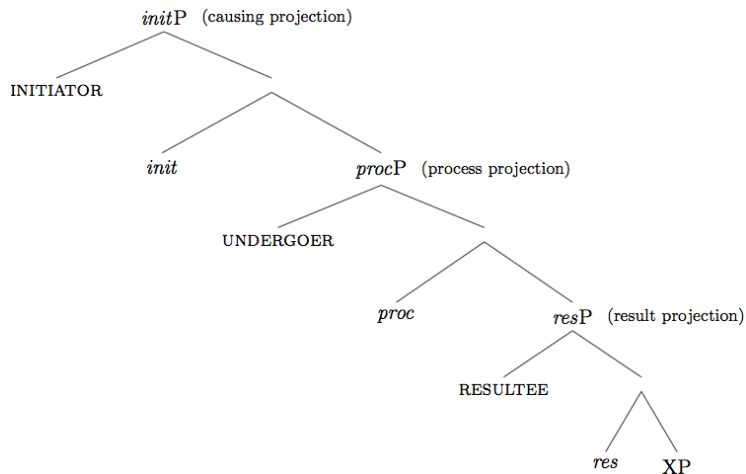
- ▶ Data basis: 65 394 matrix declarative sentences with an identifiable annotation for subject case marking from IcePaHC.
- ▶ Automatic extraction of the relevant information from these sentences via Perl scripts:
 - ▶ subject and object case marking
 - ▶ verbs and verb types (main verb, be, have, etc.)
 - ▶ auxiliaries/modals
 - ▶ voice (active, passive, middle)
 - ▶ word order, subject position, V1

ID	VERB	VERB_TYPE	MODAL/ASP	VOICE	SBJ_CASE	OBJ_CASE	OBJ2_CASE	WORD_ORDER	SUBJ_POSITION	V1
1150.FIRSTGRAMMAR.SCI-LIN,.1	setja	VB	-	active	sbj_NOM	obj1_ACC	-	VSO1	postfinite	no
1150.FIRSTGRAMMAR.SCI-LIN,.2	setja	VB	-	active	sbj_NOM	obj1_ACC	-	O1VS	postfinite	no
1150.FIRSTGRAMMAR.SCI-LIN,.3	hafa	HV	þurfa	active	sbj_NOM	obj1_ACC	-	SVO1	prefinite	no
1150.FIRSTGRAMMAR.SCI-LIN,.4	rita	VB	-	active	sbj_NOM	obj1_ACC	-	VSO1	postfinite	yes
1150.FIRSTGRAMMAR.SCI-LIN,.5	verða	RD	-	active	sbj_GEN	-	-	VS	postfinite	no
1150.FIRSTGRAMMAR.SCI-LIN,.6	ganga	VB	-	active	sbj_NOM	-	-	VS	postfinite	no
1150.FIRSTGRAMMAR.SCI-LIN,.7	rita	VB	-	active	sbj_NOM	obj1_ACC	-	VSO1	postfinite	no
1150.FIRSTGRAMMAR.SCI-LIN,.8	hafa	HV	-	active	sbj_NOM	-	-	VS	postfinite	no
1150.FIRSTGRAMMAR.SCI-LIN,.9	taka	VB	-	active	sbj_NOM	obj1_ACC	-	O1VS	postfinite	no
1150.FIRSTGRAMMAR.SCI-LIN,.10	rita	VB	-	active	sbj_NOM	obj1_ACC	obj2_DAT	VSO2O1	postfinite	no
1150.FIRSTGRAMMAR.SCI-LIN,.11	taka	VB	-	passive	sbj_NOM	-	-	VS	postfinite	no

Event decomposition: The first-phase syntax (Ramchand 2008)

- ▶ Meaning is systematically constructed as part of the syntax.
- ▶ Dynamic events can be decomposed into maximally three subevents, each with its own syntactic projection:
 - ▶ causing or initiating subevent → *initP*
 - ▶ process subevent → *procP*
 - ▶ result state → *resP*
- ▶ Each projection licenses an event participant in its specifier position:
 - ▶ causer of the event = INITIATOR
 - ▶ participant undergoing a change or process = UNDERGOER
 - ▶ participant holding the result state = RESULTEE
- ▶ A RHEME may occur in complement position of the event heads *proc* or *res* to further describe the respective subevent.

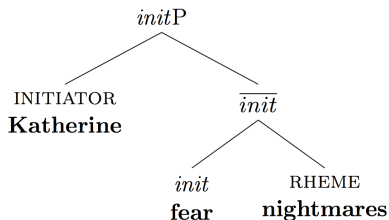
Event decomposition: The first-phase syntax (Ramchand 2008)



Event decomposition: The first-phase syntax (Ramchand 2008)

- ▶ Stative predicates do not contain a *procP*, but consist of an *init* projection only.
- ▶ With stative predicates, the INITIATOR is interpreted as a state HOLDER.
- ▶ The *init* head may furthermore select a RHEME argument which further describes the stative eventuality.

State: *x fears y*



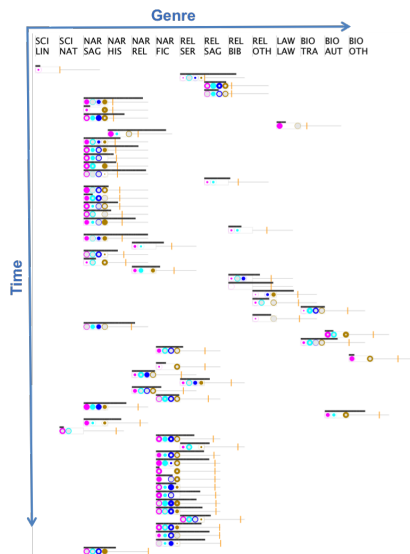
Event structure annotation

- ▶ Manual classification of dative subject predicates (main verb types) with respect to event structure
- ▶ Automatic annotation of 2004 dative subject sentences with event structure information using a Perl script

Class	Event type	Event structure	Event participants
I	Stative predicates <i>líka</i> 'like'	[<i>state</i>]	HOLDER RH HOLDER
II	Transfer <i>gefa</i> 'give'	[<i>init, proc, res</i>]	INIT RES RH INIT RES
III	Transitions <i>velta</i> 'roll'	[<i>init, proc</i>]	INIT UND UND INIT-UND RH
IV	Scalar changes <i>ljúka</i> 'end, finish'	[<i>init, proc, res</i>]	INIT UND-RES UND-RES UND-RES RH

Table 1: Event structure classification for dative subject predicates.

Glyph visualization (Butt et al. 2014, Schätzle & Sacha 2016)



- ▶ 'Overview first, zoom and filter, then details-on-demand'
- ▶ Compact representation of large amounts of data with respect to certain query terms.
- ▶ Each text from IcePaHC is visualized as glyph representation, showing different types of interaction between dative subjects, event structure, lexical semantics and voice on demand.
- ▶ Allows for a comparative diachronic evaluation of the data at different levels of detail.

Results I – Dative subjects, event structure, and voice

- ▶ The distribution of dative subjects is changing over time.
- ▶ The usage of dative subjects increases over time, with a striking increase as of 1900.
- ▶ This increase correlates with an increasing use of stative experiencer predicates together with a dative subject.
 - Mainly middle forms which have been **lexicalized** as experiencer and raising predicates with a dative subject, e.g., *finnast* 'seem'

- (1) Mér fannst það vera friður
I.DAT seem.PST.MID.3SG that.NOM be.INF peace.NOM
náttúrunar.
nature.the.GEN
'That seemed to me to be the peace of nature.'
(IcePaHC, 1920.ARIN.REL-SER,.639)

Results I – Dative subjects, event structure, and voice

- ▶ The possibility to have dative subjects correlates with particular event structural configurations.
- ▶ Dative subjects occur most frequently as state holders:
 - ▶ HOLDER with stative experiencer predicates (active, middle)
 - ▶ RESULTEE with transfer verbs (passive, middle)
- ▶ Dative case is not preserved under middle formation when the corresponding argument is an UNDERGOER(-RESULTEE).
 - ▶ transition verbs
 - ▶ verbs of scalar changes

Results I – Dative subjects, event structure, and voice

- ▶ The possibility to have dative subjects correlates with particular event structural configurations.
- ▶ Dative subjects occur most frequently as state holders:
 - ▶ HOLDER with stative experiencer predicates (active, middle)
 - ▶ RESULTEE with transfer verbs (passive, middle)→ **goals/experiencers**
- ▶ Dative case is not preserved under middle formation when the corresponding argument is an UNDERGOER(-RESULTEE).
 - ▶ transition verbs
 - ▶ verbs of scalar changes→ **theme arguments**

Subject case and word order

- ▶ Syntactic position is a main subjecthood criterion in Icelandic (see, e.g., Zaenen et al. 1985, Barðdal and Eythórssón 2003).
 - ▶ However, there are changes with respect to word order in the history of Icelandic:
 - ▶ 'freer' > less 'free' word order (Rögnauldsson 1995)
 - ▶ Germanic shift from OV to VO (Kiparsky 1996, Rögnauldsson 1996, Hróarsdóttir 2000)
 - ▶ decrease in V1 (Sigurðsson 1990, Butt et al. 2014)
 - ▶ Word order changes are a main factor behind the reanalysis of datives as subjects in historical English (Allen 1995) and other Indo-European languages (e.g., see Haspelmath 2001).
- ⇒ Investigating dative subjects in conjunction with the Oblique Subject Hypothesis necessitates an investigation of word order!

DEMO

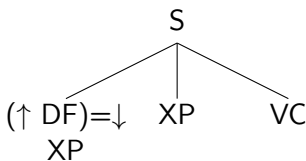
<http://subva.dbvis.de/histobankvis-v1.0/#/>

Results II – Datives, subject position and verb placement

- ▶ **Prefinite position** becomes the **preferred subject position** in the history of Icelandic.
 - ▶ V1 declarative structures decrease.
 - ▶ 19th century is a major key turning point.
 - ▶ Dative subjects lag behind in being realized in a particular position.
 - ▶ Manchester collaboration (Hannah Booth): Rise of expletives is connected to the observed changes.
- ⇒ Evidence for the development of structure and the rise of positional licensing in Icelandic.

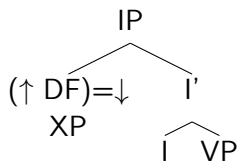
Rise of positional licensing

- ▶ Kiparsky (1995): Germanic languages developed structure and functional categories not present in Indo-European ancestor.
- ▶ Growth of structure and the development of functional categories in Icelandic noun phrases (Börjars et al. 2016).
- ▶ Early Germanic had fairly free word order, with grammatical functions indicated by case morphology.
- ▶ Flat tree in which word order is used to signal information-structural content (cf. Urdu/Hindi, Butt and King 2004).



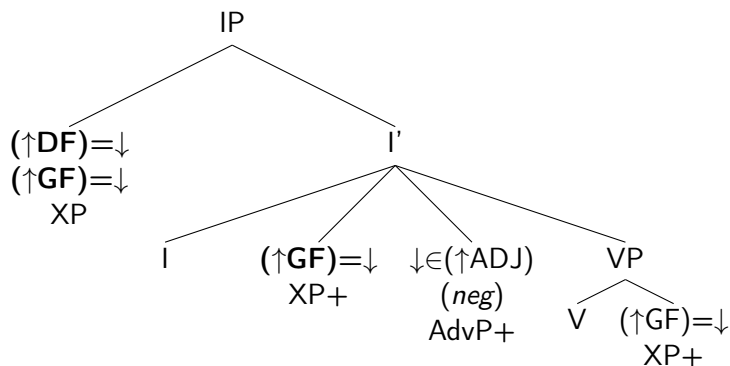
Rise of Positional Licensing

- ▶ Periphrastic tense/aspect arises, leading to an I (cf. Old English, Kiparsky 1997).
- ▶ Finite verbs (I) partition a clause in terms of information-structural information (topic vs. comment, cf. Hinterhölzl & Petrova 2010).
- ▶ V1 in topicless sentences (e.g. presentationals).



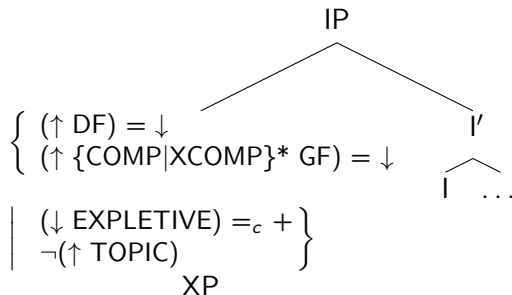
Rise of Positional Licensing

- ▶ Blueprint for clausal structure in current Icelandic proposed by Sells (2001, 2005).



Rise of Positional Licensing

- ▶ Prefinite position in SpecIP is associated with a discourse function (i.e., topic).
- ▶ Subjects tend to be topical and the SpecIP position becomes increasingly associated with subjects.
- ▶ Subjects can occur in the immediately postfinite position when the prefinite position is occupied.



Intermediate Summary

- ▶ Case marking and grammatical relations are part of a complex, yet systematic, interacting system:
 - ▶ position
 - ▶ event structure
 - ▶ lexical semantics
 - ▶ case marking
 - ▶ grammatical relations

Intermediate Summary

- ▶ Case marking and grammatical relations are part of a complex, yet systematic, interacting system:
 - ▶ position
 - ▶ event structure
 - ▶ lexical semantics
 - ▶ case marking
 - ▶ grammatical relations
- ▶ This system has been changing over the centuries.
 - ▶ Increasing systematic association between dative subjects and experiencer semantics (state holders).
 - ▶ Development of structure and rise of positional licensing
 - ▶ Dative subjects lag behind in being realized in a particular position.

Intermediate Summary

- ▶ Case marking and grammatical relations are part of a complex, yet systematic, interacting system:
 - ▶ position
 - ▶ event structure
 - ▶ lexical semantics
 - ▶ case marking
 - ▶ grammatical relations
- ▶ This system has been changing over the centuries.
 - ▶ Increasing systematic association between dative subjects and experiencer semantics (state holders).
 - ▶ Development of structure and rise of positional licensing
 - ▶ Dative subjects lag behind in being realized in a particular position.
- ▶ These results now need to be further understood.

Intermediate Summary

- ▶ Case marking and grammatical relations are part of a complex, yet systematic, interacting system:
 - ▶ position
 - ▶ event structure
 - ▶ lexical semantics
 - ▶ case marking
 - ▶ grammatical relations
- ▶ This system has been changing over the centuries.
 - ▶ Increasing systematic association between dative subjects and experiencer semantics (state holders).
 - ▶ Development of structure and rise of positional licensing
 - ▶ Dative subjects lag behind in being realized in a particular position.
- ▶ These results now need to be further understood.

⇒ **Proposal of a novel linking theory**

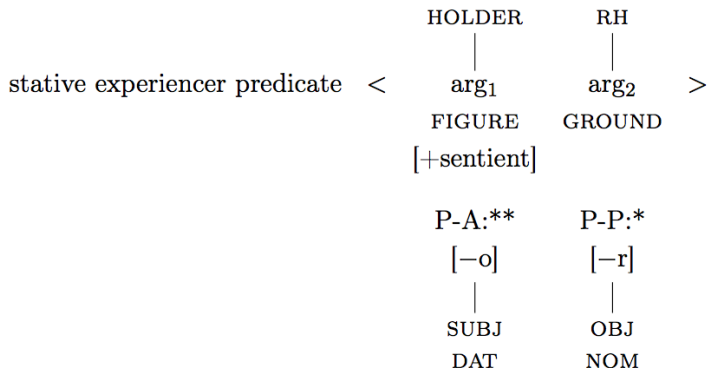
Linking events, case and grammatical relations in Icelandic

Proposal of a novel linking theory which ...

- ▶ builds on enhanced versions of LFG's Lexical Mapping Theory
 - ▶ **argument positions** (Kibort 2014)
 - ▶ **lexical semantics** (Zaenen 1993)
- ▶ incorporates a **reference frame** in the form of Talmy's (1978) FIGURE-GROUND division
- ▶ incorporates event structure via Ramchand's (2008) **event participants**
- ▶ uses Zaenen's (1993) feature classification to map **grammatical relations** to arguments.
 - ▶ Based on Dowty's (1991) Proto-Role entailments
 - ▶ Proto-Role entailments follow from the lexical semantics associated with the event participants and the reference frame
- ▶ Event structure interacts with **case marking** (in line with Svenonius 2002).

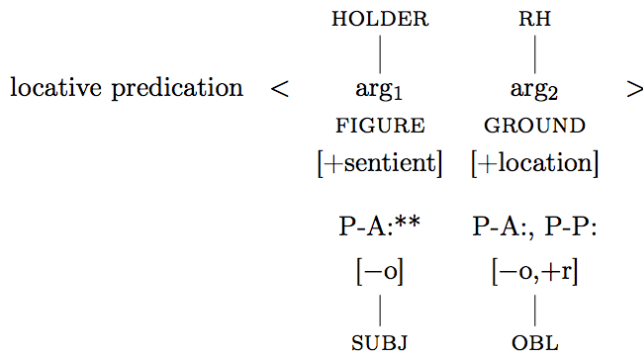
An example

- (2) En Þorgrími líkaði það illa.
but Þorgrímur.DAT like.PST.3SG this.NOM badly
'But Þorgrímur disapproved of this.'
(IcePaHC, 1400.GUNNAR.NAR-SAG,.241)



Linking analysis – Dative subjects and positional licensing

- ▶ Stative experiencer predicates with a dative subject are historically derived from locative predications.
- ▶ Locative predication: location/goal is marked dative.
→ GROUND linked to OBL

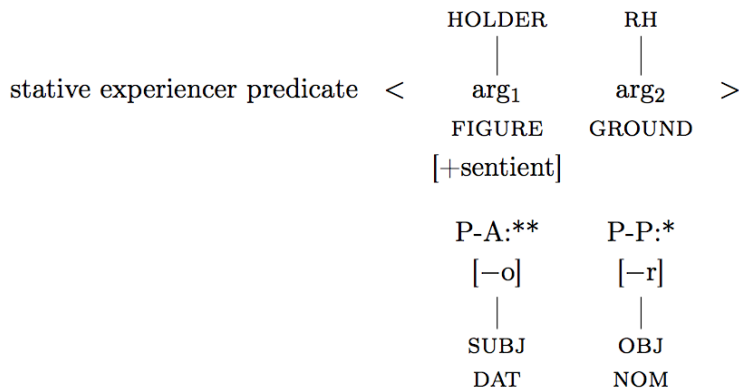


- ▶ Sentient location/goal may be interpreted as experiencer.

(3) Mikil ógn stóð mér af þessum
much.NOM fear.NOM be.situated.PST.3SG I.DAT of this.DAT
manni.
man.DAT
'A lot of fear of this man was situated **at me.**'
(IcePaHC 1300.ALEXANDER.NAR-SAG,.272)

- ▶ As sentient participants, dative experiencers are increasingly realized as state HOLDER and FIGURE arguments.
- ▶ FIGURES are usually topics and tend to occur in SpecIP.
- ▶ Over time, subjects become more firmly associated with SpecIP.
- ▶ Dative experiencers are increasingly realized as subjects.
→ HOLDER linked to SUBJ

Linking analysis – Dative subjects and positional licensing



- ▶ Dual linking possibilities for experiencers explain the weaker tendency of dative subjects to occur in a particular position.
- ▶ Lexicalized middles are instantiated as dative subjects only after they have been reanalyzed as stative predications.

Summary and conclusion

- ▶ Contribution to the on-going discussion on whether dative subjects are a common Proto-Indo European inheritance
→ Evidence for the Object-to-Subject Hypothesis
- ▶ Innovative research on Visual Analytics for historical linguistics
 - ▶ Development of novel and powerful analysis tools for historical linguistic research
 - ▶ Uncovering of previously unknown interrelations between dative subjects, lexical semantics, event structure, voice and word order in Icelandic by means of the visualizations
 - ▶ Generation of new insights with respect to the diachrony of dative subjects in Icelandic
- ▶ Development of a novel linking theory which factors in the relevant components for licensing case and grammatical relations in the history of Icelandic.

Thank you!