UNIVERSITEIT GENT

www.evalisa.ugent.be

How To Identify Cognates in Syntax: Taking Watkins' Legacy One Step Further

Jóhanna Barðdal & Thórhallur Eythórsson

erc

Overview

- The Problem: How to identify cognates in Syntax?
- Major arguments against syntactic reconstruction – and ways to deal with them
- How to identify cognates in argument structure constructions
  - On the basis of Cognate Lexical Material
  - On the basis of Cognate Structures
- Taking Watkins' Legacy one step further
  - On the basis of non-cognate synonymous material

The Problem

- Common assumption: Syntactic reconstruction is more or less untenable, because of:
    - lack of directionality in syntactic change
    - lack of arbitrariness in syntax
    - lack of continuous transmission of syntactic structures during acquisition
    - lack of simple form–meaning correspondences
    - lack of cognate material in syntax

Three Attempts in the Early Seventies

- Reconstruction of basic word order for Proto-Indo-European
  - Lehmann (1974): SOV
  - Friedrich (1975): SVO
  - Miller (1975): SOV, SVO and VSO
- Slaughtered by Watkins (1976)

Watkins (1976)

- Morphological clues are instrumental for identifying cognates, including cognate syntactic material
- Work on poetic formulae consistently identifies layers of cognate collocations and prefabs across the Indo-European traditions
- This shows how fragments of earlier syntax can be determined and hence reconstructed

Watkins (1995): How to Kill a Dragon in Indo–European

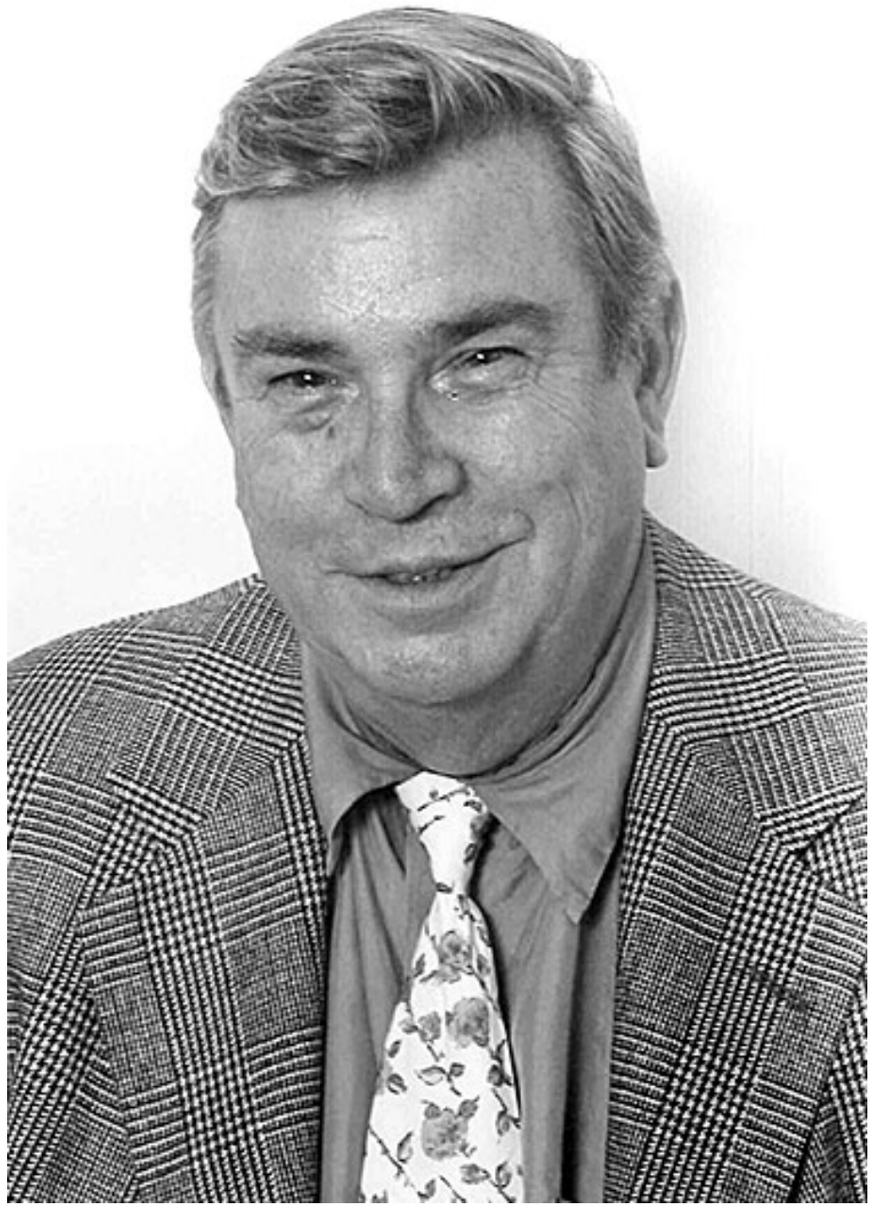
- Based on:
  - archaic expressions containing frozen syntax
  - deviations from productive patterns
  - any anomalies that cannot be explained synchronically

- One has to:
  - examine the data carefully
  - compare linguistic units used to express similar content across the daughters
  - in general, compare cognate text traditions based on oral transmission of inherited cultural and linguistic material

Taking Watkins' Method at Face Value

- Essentially all morphosyntax constitutes a potential input for correspondence sets, and therefore provides a basic for comparative reconstructions
- This is by no means an insignificant portion of grammar – *all morphosyntax*
  - In addition, collocations and prefabs may provide information about word order and clause structure of earlier stages, hence such abstract units can be detected and reconstructed.

Calvert Watkins (1933–2013)



Why Syntactic Reconstruction

- Not a hobby for armchair linguists who enjoy playing with historical data
- It is a fundamental part of historical linguistics, as it involves putting forward grounded hypotheses on pre-stages of languages, and hence how language change comes about (cf. Ferraresi & Goldback 2008)
- It thus functions as a concretization and even ”formalization” of the historical linguist’s analysis
- As such, syntactic reconstruction may provide important insights into the development of specific structures

Identifying Cognates in Syntax

- ... on the basis of cognate lexical material
- … on the basis of cognate structure, including argument- and predicate structure

- How can Watkins' program be taken one step further?
  - Via cognate recognition in syntax on the basis of synonymous lexical material

Cognate Recognition in Syntax on the Basis of Cognate Lexical Material

- One looks for lexical correspondence sets for specific verbs, e.g. Germanic ‘think’.

Lexical Correspondence Sets for Gmc 'think'

	FORM	MEANING	RECONSTRUCTED FORM
Gothic	*þugkjan*	'think, seem'	
Old High German	*thunkian*	'think, seem'	
Old English	*þyncan*	'think, seem'	**þunkjan-*
Old Saxon	*thunkian*	'think, seem'	
Old Norse-Icelandic	*þykkja*	'think, seem'	

Lexical Correspondence Sets for Gmc 'think'

Gothic

(1a) *þugkeiþ* **im** auk ei ...
thinks.3SG them.DAT because that
'for they think that ...' (Mt 6.7)

Old High German

(1b) samomichel uuunder mag **temo** *dunchen*, der ...
same.great wonder.NOM may.3SG him.DAT seem.INF who.NOM
'He will think it an equally great wonder, who ...' (Notker 1,283,9)

Old English

(1c) Ne *þynceð* **me** gerysne þæt we rondas beren eft to earde
not thinks.3SG me.DAT appropriate that we sheilds bear back to earth
'I do not find it appropriate that we bear our shields back home' (Beowulf 2653)

… Germanic 'think'

Old Saxon

(1d) than *thunkid* **imu**, that he sie gerno forð lêstien uuillie
then seems.3SG him.DAT that he it gladly forward do.INF wishes
'Then he thinks that he will gladly wish to do it in the future'
(Heliand 2496–2501)

Old Norse-Icelandic

(1e) **oss** *þykir* eigi verr að þú sért lítt heil
us.DAT seems.3SG not worse that you are little healthy
'we don't find it worse that you are not well' (Fóstbræðra s., Ch. 10

Predicate-Specific Correspondence Sets for the Argument Structure of Germanic 'think'

	ALT 1	ALT 2	ALT 3
Gothic	DAT-'thinks'		
Old High German	DAT-'thinks'	ACC-'thinks'	
Old English	DAT-'thinks'		
Old Saxon	DAT-'thinks'		
Old Norse-Icelandic	DAT-'thinks'		
Middle High German	DAT-'thinks'	ACC-'thinks'	NOM-'think'
Middle English	DAT-'thinks'	ACC-'thinks'	
Middle Dutch		ACC-'thinks'	
Old Swedish	DAT-'thinks'	ACC-'thinks'	
Modern High German	DAT-'thinks'	ACC-'thinks'	NOM-'think'
Modern Icelandic	DAT-'thinks'		
Modern Faroese	DAT-'thinks'		
Modern Swedish			NOM-'think'
Modern English			NOM-'think'
Modern Dutch	DAT-'thinks'	ACC-'thinks'	

Reconstruction of the Argument Structure of 'think' for Proto-Germanic, containing a Dative Subject

- The earliest representatitives of Germanic have a dative subject
- Accusative and nominative subjects are an innovation, attested first in later texts, also in accordance with known developmental paths of oblique subjects

Reconstruction of ’think’ for Proto-Germanic

- This proposal amounts to claiming not only that the predicte itself is cognate, but also its argument structure

- We have identified the argument structure as a cognate argument structure, inherited from a common proto-stage, on three grounds:
  - The lexical predicate (including both its form and meaning) is cognate
  - The case frame itself is cognate
  - The morphological case markers are cognate

Reconstruction of the Argument Structure of 'think' in Proto-Germanic

lxm		
FORM	< *þunkjan >	
SYN	ARG-ST <NP-DAT$_i$>	
SEM	FRAMES	regard-fr COGNIZER i

Lexical Correspondence Sets for Gmc 'thirst'

	FORM	MEANING	RECONSTRUCTED FORM
Gothic	*þaursjan*	'thirst'	
Old High German	*dursten*	'thirst'	
Old English	*þyrstan*	'thirst'	**þurst-*
Old Saxon	*thurstian*	'thirst'	
Old Norse-Icelandic	*þyrsta*	'thirst'	

Predicate-Specific Correspondence Sets for the Argument Structure of Germanic 'thirst'

	Alt. 1	Alt. 2	Alt. 3
Gothic	ACC-'thirsts'		
Old High German	ACC-'thirsts'		
Old English	ACC-'thirsts'	DAT-'thirsts'	NOM-'thirst'
Old Saxon	ACC-'thirsts'		
Old Norse-Icelandic	ACC-'thirsts'		
Middle High German	ACC-'thirsts'		
Middle English	ACC-'thirsts'	DAT-'thirsts'	
Middle Dutch	ACC-'thirsts'	DAT-'thirsts'	
Old Swedish	ACC-'thirsts'		
Modern High German	ACC-'thirsts'		NOM-'thirst'
Modern Icelandic	ACC-'thirsts'		
Modern Faroese			NOM-'thirst'
Modern Swedish			NOM-'thirst'
Modern English			NOM-'thirst'
Modern Dutch			NOM-'thirst'

Three Main Arguments

- … for assuming that the accusative subject construction is inherited from a common proto-stage:
  - the lexical predicate (including both its form and meaning) is cognate
  - the case frame itself is cognate
  - the morphological case markers are cognate

Reconstruction of the Argument Structure of 'thirst' in Proto-Germanic

lxm		
FORM	< *þurs(t)jan >	
SYN	ARG-ST <NP-ACC$_i$>	
SEM	FRAMES	need-for-intake fr NEEDER i

One more example: Nom-Dat Construction

- Two verbs in the early layers that mean ‘answer’ and both take a Nom-Dat argument structure construction
  - **(and)swaran*
  - **(and)wurdjan*
- How do we know which one is older?

Lexical Correspondence Sets and Reconstruction for Gmc **(and)swaran*

	FORM	MEANING	RECONSTRUCTED FORM
Old English	*andswarian*	'answer'	
Old Frisian	*andswara*	'answer'	
Old Norse-Icelandic	*svara*	'answer'	*(*and*)*swaran*
Old Swedish	*svara*	'answer'	
Old Danish	*swaræ*	'answer'	

Lexical Correspondence Sets and Reconstruction for Gmc **(and)wurdjan*

	FORM	MEANING	RECONSTRUCTED FORM
Gothic	*andwaurdjan*	'answer'	
Old High German	*antwurten*	'answer'	
Old Frisian	*andwerda*	'answer'	**andwurdjan*
Old Saxon	*andwurdian*	'answer'	
Old English	*andwyrdan*	'answer'	
Middle Dutch	*antwerden*	'answer'	

Which One is Older?

- *(*and)svaran* is confined to North-Germanic and Ingveonic
- *(*and)wurdjan* is confined to East and West Germanic

- All three branches have *sverja* 'swear (an oath)' which is derived from *svara*

- Where does the case frame of the two come from?
  - *(and)wurdjan comes from the simple verb **wurdjan*, which is denominal from the noun **wurdan*. In the languages that have the verb 'word', it takes a Nom-Acc

A Reconstruction of the Argument Structure of 'answer' in PGmc

lxm

FORM	< *(and)swaran >	
SYN	ARG-ST <NP-NOM$_i$, NP-DAT$_j$>	
SEM	FRAMES	communication_response fr SPEAKER i ADDRESSEE j

Summary on Germanic 'answer'

- **(and)swaran* is the older verb in Germanic meaning 'answer', and was replaced by **andwurdjan* in East and West Germanic.
- This raises the question of from where the Nom-Dat case frame of **andwurdjan* comes. Since **andwurdjan* is derived from a noun, the case frame cannot be inherited from the source. The languages that have the verb 'word', show a Nom-Acc case frame.
- The most likely source of the case frame of **andwurdjan* is that it comes from its synonymous verb **(and)swaran*.
- In other words, when **andwurdjan* changed its meaning from 'put into words, speak' to 'answer' it also took over the case frame of **(and)swaran*.

Argument Structure Constructions with New Verbs, attracting case frames from already existing synonymous verbs

Cognate Verb Source		Synonymous Verb Source	
stoða 'support'	*aðstoða* 'assist	*bjarga* 'save'	*redda* 'save'
senda 'send'	*símsenda* 'fax'	*laga* 'fix'	*ordna* 'fix'
rita 'write'	*afrita* 'copy'	*hringja* 'phone'	*bjalla* 'phone'
færa 'move'	*uppfæra* 'update'	*daga* 'dawn'	*dona* 'dawn'
þýða 'translate'	*bakþýða* 'decompile	*skera* 'cut'	*kötta* 'cut'
baka 'bake'	*afbaka* 'distort'	*binda* 'tie'	*leisa* 'lace'
skoða 'observe'	*endurskoða* 'audit'	*hreinsa* 'cleanse'	*sjæna* 'make shiny'
skipta 'divide'	*lagskipta* 'stratify'	*trufla* 'bother'	*bögga* 'bug'
setja 'put'	gróðursetja 'plant'	*eyða* 'delete'	*dílíta* 'delete'

Identification on the Basis of Cognate Structure

- A device for identifying cognate argument structure constructions is through the Double Cognacy Requirement (Walkden 2009)
- For example /p/ cannot be reconstructed for Proto-Indo-European unless it is found in cognate enviroment
- Thus, one cannot reconstruct /p/ on the basis of *piscis* ‘fish’ in Latin and *fadar* ‘father’ in Gothic
- But on the basis of *piscis* ‘fish’ in Latin and *fisks* ‘fish’ in Gothic and/or on the basis of *pater* ‘father’ in Latin and *fadar* ‘father’ in Gothic
    - A secure reconstruction is carried out on the basis of a cognate context and not above cognate contexts.

Double – or even Triple – Cognacy Requirement

- Lexical material, predicate structure, case frame

Gothic

(5a) *goþ ist* **unsis** her wisan
good is.3SG us.DAT here be.INF
'it is good for us to be here' (Luke 9: 33)

Old English

(5b) *betere is* manna **gehwylcum** þæt him her on worulde
better is.3SG men.GEN each.DAT that him here on world
'Every man has it better here in the world' (Ælfric Homily 28.107)

Old Norse-Icelandic

(5c) *Betra er* **lifðum** **og** **sællifðum**.
better is.3SG the.lived.DAT and the.well.lived.DAT
'Those who live and live well have it better.' (Hávámál 70)

Applying this on Regular Structures, like the Nom-Acc

Gothic	Old English	Old Norse-Icelandic	Gloss
dreiban	*drīfan*	*drífa*	‘drive’
maurþrjan	*myrðran*	*myrða*	‘murder’
slahan	*slean*	*slá*	‘hit’
saggjan	*senkan*	*sökkva*	‘sink’
bairan	*beran*	*bera*	‘carry’
drigkan	*drincan*	*drekka*	‘drink’
itan	*etan*	*eta*	‘eat’
saihwan	*sēon*	*sjá*	‘see’
finþan	*findan*	*finna*	‘find’

Applying this on Regular Structures, like the Nom-Acc

- Do we also have Double Cognacy here?
- Default pattern
- Typologically well-attested
- Not necessarily any arbitrariness

- However, the morphological case markers are cognates
- The argument structure is cognate
- The lexical predicates are cognate

Identification on the Basis of Synonymous Material

- Cognate argument structure construtions may be identified despite a deeper time span
    - Cognate argument structure constructions may be distinguished, not only because they are lexical cognates, but also by virtue of being instantiated by synonymous verbs
- Since lexical predicates tend to be replaced with synonymous predicates through time, the argument structure constructions may remain stable

The Argument Structure of 'answer' in Indo-European

Ancient Greek (Attic): *apokrinomai* 'reply, answer'

(6a) egō gar autik' apokrinoumai soi saphōs
I for right.away will.answer.1SG you.DAT.SG clearly
'for I will presently answer you distinctly.' (Aristophanes, Clouds 1245)

Latin: *respondeo* 'reply, answer'

(6c) legatis respondit diem se ad deliberandum sumpturum
ambassadors.DAT answered day.ACC self to deliberate take
'He [Caesar] replied to the ambassadors, that he would take time to deliberate' (Caesar, Gallic War 1.7)

Gothic: *andwaurdjan* 'answer'

(6d) þu hwas is, ei andwaurdjais guda?
you who are that reply.OPT God.DAT
'Who are you replying to (OR: against) God?' (Rom 9:20)

The Argument Structure of ’answer’ in Indo-European

Old Russian: *otvéčal-* ‘answer’

(6e) I Pskovъ **imъ** *otvéčalъ*
and Pskov.NOM them.DAT replied.3SG
‘And Pskov answered them.’ (Pskovskaja letopis’ XVc)

Old Lithuanian: *atsakyti* ‘answer’

(6f) Ar šitaipo **byskupui** *atsisakai*?
do this bishop.DAT answer.2SG
‘Do you answer the bishop in this manner?’
(Bretkunas Postille I372, y. 1591)

Hittite: *āppa mema/i-* ‘speak/say back’

(6g) ^{d}UTU-uš *ANA* MUNUS.LUGAL *āppa memišta*
sun.god-NOM DAT queen back spoke.3SG
‘ The Sun God replied to the queen.’ (KBo 20.82 ii 33–34)

The Argument Structure of ’answer’ in Indo-European

Tocharian B

(6h) *wätk-* ‘answer’

kupre ne säm penu sne täṅklune wätkālts wätkāṣṣ-äm ////
if he PTCL without.difficulty with.confidence
answer:3SG.CONJ –CL.OBL
‘If he responds to you without difficulty and with confidence’ (YQ-14[II.5] b4)

Sanskrit: *prati-brū-* ‘answer’

(6i) apṛccha**ṃ** mātaram \ sā **mā** *pratyabravīt*
asked.3SG mother.ACC she.NOM me.ACC answered.3SG
‘I asked my mother and she answered me’ (Ch. 4.4.4)

Predicate-Specific Correspondence Set for the Argument Structure of PIE 'answer'

	Alt. 1.	Alt. 2
Ancient Greek	Nom–Dat	
Latin	Nom–Dat	
Gothic	Nom–Dat	
Old Russian	Nom–Dat	
Old Lithuanian	Nom–Dat	
Hittite	Nom–Dat	
Tocharian B	Nom–Dat	Nom-Acc
Sanskrit		Nom–Acc

Reconstruction of the Argument Structure of PIE 'answer'

lxm

FORM	< *'answer' >	
SYN	ARG-ST <NP-NOM$_i$, NP-DAT$_j$>	
SEM	FRAMES	communication_response fr SPEAKER i ADDRESSEE j

Conclusion (1)

- Watkins' work on poetic formulae shows how layers of cognate collocations and prefabs can be identified through morphological clues, together with important fragments of syntax from earlier periods of the Indo-European languages, reconstructable as such.
- We have demonstrated how cognate argument structure constructions may be identified, with the aid of
  - the lexical predicates that instantiate them
  - cognate case frames
  - cognate predicate structures, and
  - cognate case morphology

Conclusion (2)

- For this purpose, we have compared case and argument structure constructions of Germanic verbs, e.g. ‘think’, ‘thirst’ and ‘answer’ which all have case frames that deviate from the canonical Nom-Acc frame, and hence they exhibit a certain degree of arbitrariness
- The directionality of the changes is retrievable from documented processes and the dataset themselves

Taking Watkins' Research Program one Step Further

- Cognate argument structure constructions may be identified on the basis of noncognate synonymous predicates.
- This claim is based on documented processes of how new verbs acquire their case and argument structure constructions, of which one major process involves new verbs attracting case frames from already existing synonymous verbs.
- Cognate argument structure constructions may be identified using noncognate lexical material through a case study of the verb 'answer' which has two cognate sets in Germanic, but at least eight sets across Indo-European.
- This larger cognate set is expected given the greater time depth for Proto-Indo-European than for Proto-Germanic, and given current knowledge of the speed of lexical replacement.

Ultimately …

- Semantic spaces can be reconstructed for argument structure constructions at earlier proto-stages, on the basis of lexical-semantic verb classes, since argument structure constructions may remain stable while lexical items are replaced.

"That's all Folks!"