Valence: Towards a multilingual setting

Based on the paper: *Hellan et. al 2014. MultiVal – towards a multilingual valence lexicon, LREC 2014.*

# Plan

- MultiVal: current (For 2013 – see http://moin.delph-in.net/SaarlandValence)
- MultiVal: functionalities
- MultiVal: search interface
- MultiVal: from lexicons to verb valence typology
- MultiVal: extendability

MultiVal: current

- valence lexicon derived from lexicons of computational HPSG grammars for Norwegian, Spanish and Ga

- 22,000 verb entries and on average more than 200 valence types defined for each language

- mapped onto a common set of discriminants with a common array of values, and stored in a relational database linked to a web demo and a wiki presentation

MultiVal: functionalities/discriminants

- SAS (syntactic argument structure) – NP+NP; NP+PP, etc.

- FCT (functional label) – transitive, intransitive, ditransitive, etc.

- SIT* (situation type, -arity)

- ASP (aspect, Aktionsart)

Word order neutral
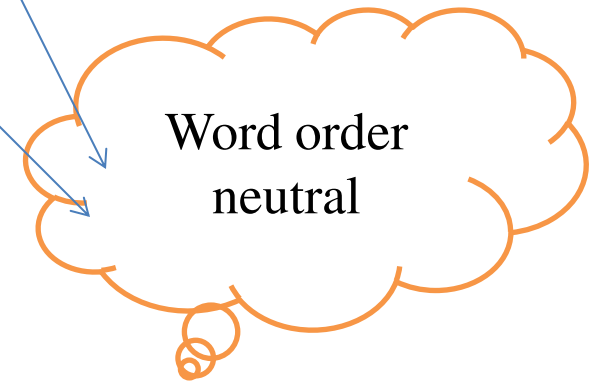
MultiVal: examples (NOR)

v-intr
SAS: "NP"
FCT: intransitive
SIT:
Example:
 gutten hopper
 gutten sover
 gutten lytter
 isen smelter
 gutten fryser

MultiVal: examples (NOR)

v-intrObl-NONCOMPLETED_MONODEVMNT
SAS: "NP+PP"
FCT: intransWithOblique
ASP: noncompleted
SIT: binaryRel
Example:
 Ola tygger på eplet
 hun spiser av eplet
 hun leser i boken

MultiVal: examples (SPANISH)

v_-_nsbj_le

SAS: +

FCT: intransImpers

SIT: weatherProcess

Example:

 llueve

MultiVal: examples (SPANISH)

v_cp_q_le

SAS: NP+Squest

FCT: transWithSentCompl

SIT:

Example:

 dijo quién ganaría

MultiVal: search interface

Multilanguage Valency Patterns

Version 1.2 (for further guidelines, see Info)

Languages:

☑ Norwegian ☑ Ga ☑ Spanish

Search fields:

V-key

Syntactic Arguments

Function   Situation   Aspect   Type

Search Count Clear Download

Search: overview

- Each language employs only a subset of the total set of values for each discriminant

- Any combination of discriminants can be applied

- The search can be done only by verb

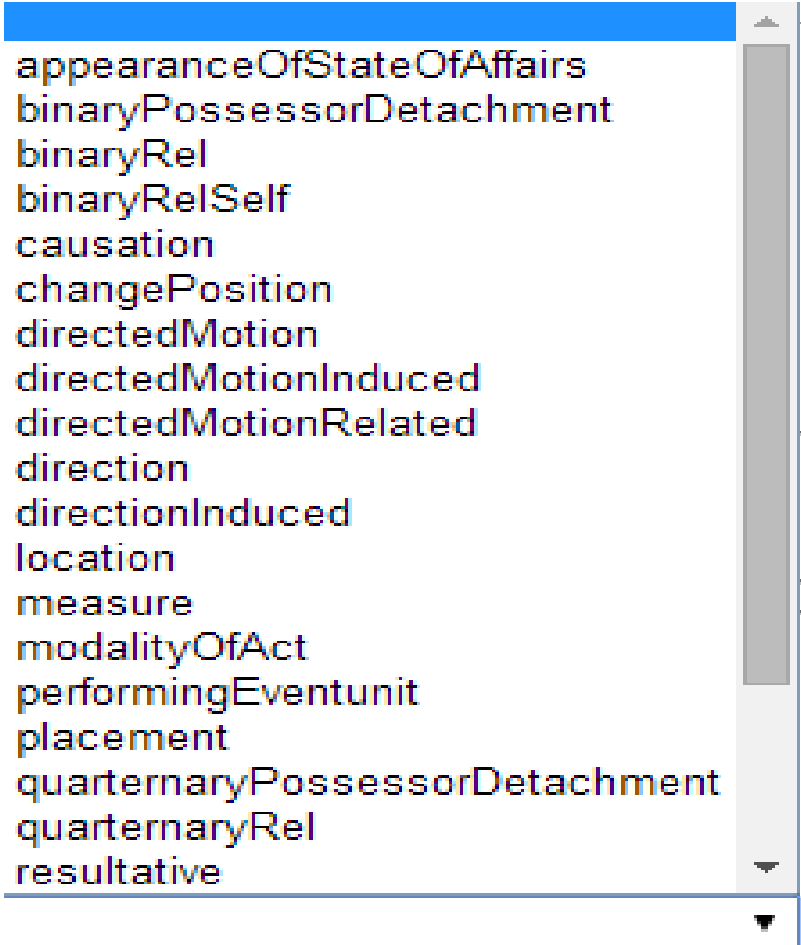- The result is the list of carrying verbs that conform to the query

Syntactic Arguments

```
+
+INF
+NP
+PP
+PP[INF]
+PP[Sdecl]
+PPa+PP
+PPa+PP[Sdecl]
+Sdecl
EXPL+APpred+S
EXPL+APpred+adpos
EXPL+INF
EXPL+NP
EXPL+NP+INF
EXPL+NP+INF[equiOBJ]
EXPL+NP+NP+INF
EXPL+NP+Sdecl
EXPL+NP+SquestWH
EXPL+NP+SquestYN
```

Function

copulaImpersonalWithPredicativeAdverbAndLocative
copulaWithIdentityDeclClause
copulaWithIdentityInfinitive
copulaWithIdentityNoun
copulaWithIdentityWhClause
copulaWithIdentityYesNoClause
copulaWithPredicativeAdjective
copulaWithPredicativeAdjectiveAndExtraposedClause
copulaWithPredicativeAdjectiveAndSententialSubj
copulaWithPredicativeAdv
copulaWithPredicativeAdverbAndSententialSubj
copulaWithPredicativeNoun
copulaWithPredicativeNounAndExtraposedClause
copulaWithPredicativePP
copulaWithPredicativePPAndExtraposedClause
copulaWithPredicativeParticlephrase
diransWithSentComplAndOptIndObj
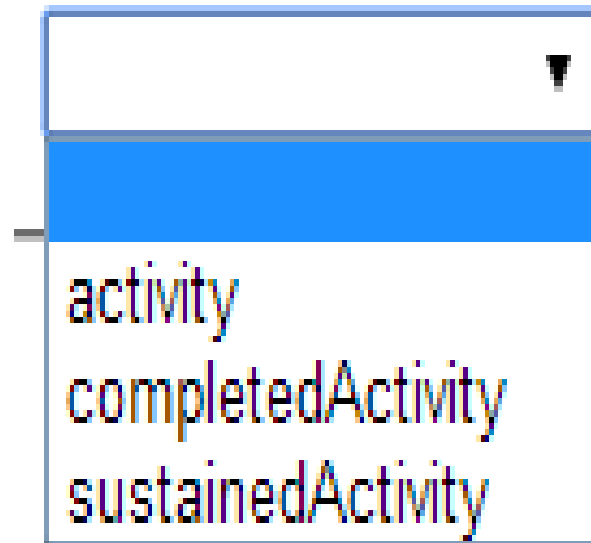ditransLightSecobj
ditransReflexWithOptIndObj

Situation

- appearanceOfStateOfAffairs
- binaryPossessorDetachment
- binaryRel
- binaryRelSelf
- causation
- changePosition
- directedMotion
- directedMotionInduced
- directedMotionRelated
- direction
- directionInduced
- location
- measure
- modalityOfAct
- performingEventunit
- placement
- quarternaryPossessorDetachment
- quarternaryRel
- resultative

Aspect

Aspect

▼

activity
completedActivity
sustainedActivity

Type

v-ditr-suINTERR_obINTERR
v-ditrComp
v-ditrExpnSu-obMeas_expnEqInf
v-ditrLght
v-ditrObl-oblPRTOFiob-PARTWHOLE_AFFECTING
v-ditrVid
v-intr
v-intr-suAbsinf
v-intr-suDECL
v-intr-suDir
v-intr-suDir-suMover
v-intr-suINTERR
v-intrAdj
v-intrAdv
v-intrAuxmodComp-compEqBareinf
v-intrAuxmodScpr-scSuNrg_scBareinf
v-intrAuxpassScpr-scSuNrg_scPass
v-intrAuxperfScpr-scSuNrg_scPerf

MultiVal: from lexicons to verb valence typology - overview

	NorSource	SRG	GaGram
Verb types	348	236	144
Entries	12500	8000	2000
Optional Arguments	Has an entry for each separate frame	Subsumes transitive and intransitive usages in one	Has an entry for each separate frame

MultiVal: from lexicons to verb valence typology – the database

- Copying information from lexical entries

- Manually created conversion list for each language from specific lexical types to MultiVal discriminants.

Currents Issues: Differences and Gaps

- Granularity of the encoded information into the specific language lexicons (syntactic or also semantic)
- Specifications followed in the lexical entries and type labels (previous undertakings, other grammars, etc.)
- Levels of mapping into the MultiVal resource (SAS and FCT filled, but SIT and ASP – very sporadically filled)
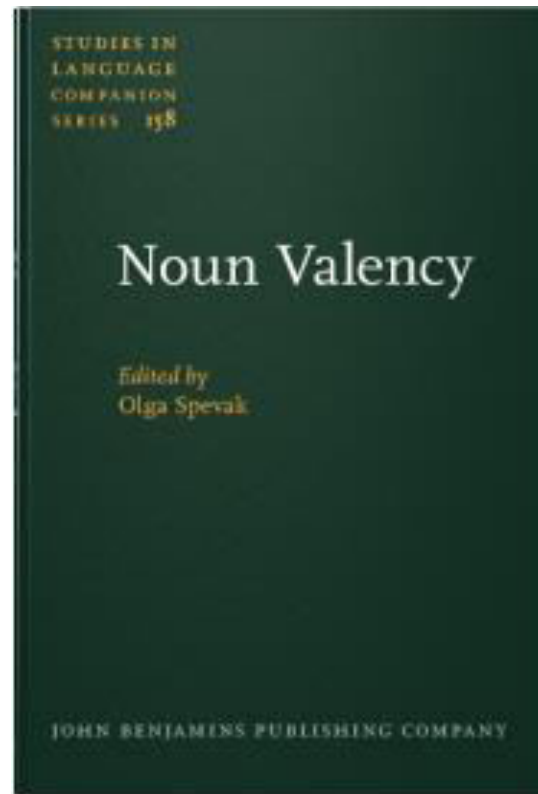- Glossing in English

MultiVal: extendability with more languages

Bulgarian

Specifics:

- The current types manually mapped

- at least 200+ fine-grained verb entries and about 18 000 coarse-grained

- SIT not filled

- ASP filled as perfective/imperfective

- Glossed in English only for the verbs

MultiVal: extendability with more valence-taking POS

MultiVal: extendability via adaptation of the current discriminant schema

- Nouns (deverbal, relational, quantity, etc.)

SAS, FCT are always relevant

SIT might be relevant (Qualia Structure??)

ASP might not be relevant

- Adjectives

- Prepositions (ASP excluded?)

The reduction of the discriminants depends on the interpretation

Sum up

- Are the presented discriminants for verb valency universal enough?

- What about the dependency on the specific grammar? Does the current setting perform the mapping to discriminants OK?

- Can we adapt the schema to other valency-sensitive POS?