# Report for PARSEME Short Term Mission, reference code: COST-STSM-ECOST-STSM-IC1207-230315-055491. by Lars Hellan, NTNU, and Petya Osenova, IICT-BAS

Trondhein/Sofia 21.04.2015

From 23 March to 30 March the mission took place at IICT-BAS, Sofia, with the grantee Lars Hellan as visitor and Petya Osenova and Kiril Simov as hosts, in pursuing the objectives stated in the mission proposal:

The aim of the mission is to enhance the Norwegian-Bulgarian valence lexicons, used in the language deep grammars, with MWEs of various types. Special focus will be put on the following linguistic features: the degrees of compositionality and fixedness. As results we expect: A meaningful comparison of MWEs in the two languages and enrichment of the lexicons with new lexical entries, which would approve the accuracy of the grammars, and enhance the expressiveness and accessibility of the verb information in the valency databases of the languages.

In addition to intensive working sessions every day, Prof. Hellan also gave two presentations, one in the institute (IICT) and one for the BulTreeBank project group.

#### Background

We had at our disposal the following prerequisites: the *Norsource*<sup>1</sup> and *BURGER*<sup>2</sup> *HPSG*<sup>3</sup> grammars; valency lexicons as part of the grammars; the *BulTreeBank*<sup>4</sup>, the annotation platform *TypeCraft*<sup>5</sup>; the online multilingual valency lexicon *MultiVal*<sup>6</sup>; a partial inventory of MWE types in each language; valency frames from BulTreeBank and from Norsource. We had decided to focus on verb-headed MWEs in both languages, and among them, on light verb constructions and semi-fixed idioms, since Bulgarian lacks phrasal verbs. Our approach would include the idea of using *catena*<sup>7</sup> in the encoding of MWEs in the lexicon and in the data.

#### Results

1. An appropriate encoding of MWEs in the TypeCraft system.

The chosen encoding is illustrated below, for the Norwegian MWE *stå på agendaen* ('be on the agenda'; see (2a) below for a more detailed view of the TypeCraft annotation interface):

(1)	stå	på	agendaen
	HEAD.CAT1	OBL.CAT1	POBJ.CAT1
	stå	on	agenda-the

The encoding marks the parts of the catena (CAT1) over grammatical roles (HEAD, OBL(ique), POBJ (prepositional object)). Annotated examples including this information will be used for enriching the lexicons in grammars and in the MultiVal lexicon.

#### 2. Annotation of 30 parallel (BG-NO) sentences containing MWEs in TypeCraft.

The annotation includes: catena over the MWEs, free translation into English and morphosyntactic information in the form of interlinear glossing. The English translation constitutes the common

<sup>1</sup> Norsource: <u>http://typecraft.org/tc2wiki/Norwegian HPSG grammar NorSource</u>; Beermann and Hellan 2004, Hellan 2008.

<sup>2</sup> See Osenova 2011.

<sup>4</sup> See Simov et. al 2005

<sup>&</sup>lt;sup>3</sup> For introduction to the general framework, see Pollard and Sag 1994, and for the implementation format of the grammars used, the LKB platform, see Copestake 2002. The platform is generally maintained through the DELPH-IN consortium, cf. <u>http://moin.delph-in.net/</u>

<sup>&</sup>lt;sup>5</sup> <u>http://typecraft.org/tc2wiki/Main\_Page</u>; Beermann and Mihaylov 2013.

<sup>&</sup>lt;sup>6</sup> <u>http://regdili.hf.ntnu.no:8081/multilanguage\_valence\_demo/multivalence</u>; Hellan et al. 2014.

<sup>&</sup>lt;sup>7</sup> See Simov and Osenova 2014

denominator for the corresponding Bulgarian and Norwegian sentences. Beneath are screenshots illustrating in (2a) the Norwegian member of such a pair, with MWE annotation following the format shown in (1), and in (2b) the result for a search leading to this pair, prompted by its common meaning, as exhibited in the search interface in TypeCraft:

		oplications Places Sys	stem 😒 🎒 🔗	<u>a</u>		-10	2 3	2	1 1 1 1 1 1 1	21.	4			* 🖘 🖬				A Ihelian 🖒
		🛇 🕤 TC Editor - M	lozilla Firefox															
the specified of t								a de la companya de l	en dia							-		
Prime * month   ************************************					Problem	10	My Texts	[]] TC	Editor 1	NorS	burce SIR Ho	ourly fore	Malo				a Leo	
a. best fib i				2061,54740											typecra	rc.org		4
			and i															
Image: registration       Image: registration<		Save Share: B		Publish	Templat	te h	ew phrase	e Dele	te phrase	View p	hrase list	View disc	ourse sense	es				
Image: set is a backet for setting:   Image: set is a backet for set is a	Provide and the state in the state in the state in the state.   Provide and the state in the state in the state in the state.   Provide and the state in the state in the state in the state.   Provide and the state in the state in the state in the state.   Provide and the state in the state in the state in the state.   Provide and the state in the state in the state in the state.   Provide and the state in the state in the state in the state.   Provide and the state in the state in the state in the state.   Provide and the state in the state in the state.   Provide and the state in the state in the state.   Provide and the state in the state in the state.   Provide and the state in the state in the state.   Provide and the state in the state in the state.   Provide and the state in the state in the state.   Provide and the state in the state in the state.   Provide and the state in the state in the state.   Provide and the state in the state in the state.   Provide and the state in the state in the state.   Provide and the state in the state in the state.   Provide and the state in the state in the state.   Provide and the state in the state.   Provide and the state.	16. Du har ingen r	ret *														×	
Image: set of a load of a region	Prevention   Prevention <td>B Save FTr</td> <td>ans 1 FTrans</td> <td>2 CParan</td> <td>n Base</td> <td>Mean</td> <td>ing Glo</td> <td>ss POS</td> <td></td>	B Save FTr	ans 1 FTrans	2 CParan	n Base	Mean	ing Glo	ss POS										
Image: set of a set	The standard is The standard is         (a)	Phrase: V	Du har inge	n rett til å h	olde meg a	nsvarlig fo	or fortiden.	-										
	Where Comparison   Where	Free translation 1:		ave the rig	nt to hold n	ne respon:	sible for th	ie past										
	Image: provide a state of the sta	The during out 2.		NR-NR-AR	rod													corpus
						rett	rit	A	bolde	mee	ansvartie	for	fortideo		1.1		140	1 Expressions. Re
	Search frame:       Image:	and the second s		10.12	00000		1000	100		14441723	and constant of the		1000 M 1000 M	en				
Several term is ab 2.50.400 km per begin due to the per begin	Image: Data State S	Baseform:		ha r	Ingen		til	å	1000000	jeg			fortid	en				
Propried   Propr	Image by a galary data data water patient in the galary			AT2 PDFS	100.777.0		1.1.1	INF CATO		150 400		1000 C	past	DEESG MASC				
Image: state in the state	Image: state   Image: state <td></td> <td></td> <td>v</td> <td></td> <td></td> <td>177 10 T</td> <td></td> <td></td> <td></td> <td></td> <td></td> <td>N</td> <td>Provide Commence</td> <td></td> <td></td> <td></td> <td></td>			v			177 10 T						N	Provide Commence				
Add dictores area   Interminity    Interminity    I	Add decurrent enter   Declared and the server in the serv										0							
Visite and the set of	Cited on the distance of the distan	C2						A.									-	
Intersent plane gave appendix on the set and before the set	In proper day any paper day any p		ense														1	
be denote subset of the subset	bit dense may call for dense as fuences of find a scouter for hold some are reporting in the source in report of the source in the source in report of the source in the sourc		ин държнш смет	а за минал	010.												-11	
Advance hand is earbeidded - schusily 2 control cont   Tor Editor - Modella Precise   Tor Editor - Modella Precise   Reclatories Places System () () () () () () () () () () () () ()	leivent haad is enterdeded - extually 2 gegeges C2 ha egg 31 have right rot to right of the rest with a holde meg ansvariig for fortiden. To Extor: Montila Fieror:	holde neen ansvar	rlig for - (държа н	метка за [м	нналото] (	'hold acco	ount for' he	old someon	e responsi	(ble)							4	
Tt Editor - Maxilla Firefox         Including Flaces System Signal Firefox         Search Phrase - TypeCraft - Moxilla Firefox         Search Phrase         Search P	tenden eret til å hade meg ansvallig for fortiden.					ett til hav	e right to											
Tt Editor - Maxilla Firefox         Including Flaces System Signal Firefox         Search Phrase - TypeCraft - Moxilla Firefox         Search Phrase         Search P	The definition functions system is a responsible definition of the past in the																1	
Aladorius Places System       Image: Control Places System       Image: Control Places System       Image: Control Places       I	Alactions Places System Since Holes       Image Holes       Image Holes																	
Aladorius Places System       Image: Control Places System       Image: Control Places System       Image: Control Places       I	dictions Places System       Image: Ima	har ingen rett til å	holde meg an	svarlig for	fortiden								¥	E+		1		8
• y typecraft.orgit.2wikl/Special/TypeCraft/SearchPhrase/           • e e e e e e e e e e e e e e e	typecaft.org/t.2wiki/Special:typeCraft/Spe	) TC Editor - Mozilla F pplications Places Sy	irefox stem SØ?					5	T T		4		¥			Q.	۲	€ Ihelian (
Lars Hellan my talk my preferences my watchlist my contributions log     arrer 4 sitt arrer few 1 sitt arrer few 1 sitt arrer me be peaced along in the forest?"     arrer 4 sitt arrer few 1 sitt arrer few	Lars Helian my taik my preferences my watchilst my contributions by strate if uftrations for a data of the ground land of the	) TC Editor - Mozilla F plications Places Sy Search Phras Edit View History E	irefox stem SØ ? se - TypeCraft Bookmarks <u>T</u> ool	) ← Mozilla s <u>H</u> elp	Firefox				-		¢.			: ∦ ╤ ■)		¥.	<b>V</b>	
Translation       Contributor       Last changed books and the past         St search Phrase       Phrase       Translation         Bokmåli       Du har ingen rett il å holde meg ansvarlig for for the past.       You do not have the right to hold me responsible for the past.         St sags list ene tags list ene	With Bitly	) TC Editor - Mozilla F plications Places Sy C Search Phras Edit View Higtory E avano - n	irefox stem S 2 ? se - TypeCraft Bookmarks Tool Act D TC I	o <b>∽</b> - Mozilla s <u>H</u> elp cditor	Firefox				C Editor	) () Nor	Source Surte	Hourly for		\$ 중 •■ slGram			🔮 nna Leo.	🔲 Multilani
ininglal meth m	Sectal page                  Sectal page <td>) TC Editor - Mozilla F plications Places Sy</td> <td>irefox stem S 2 ? se - TypeCraft Bookmarks Tool Act D TC I</td> <td>o<b>∽</b> - Mozilla s <u>H</u>elp cditor</td> <td>Firefox</td> <td></td> <td></td> <td></td> <td>ू प्रि</td> <td>Nor:</td> <td></td> <td></td> <td>е 🦳 Ма</td> <td>* ♀ ■ slGram</td> <td>▼ typec</td> <td>raft.org</td> <td></td> <td> ) 📰 Multilan</td>	) TC Editor - Mozilla F plications Places Sy	irefox stem S 2 ? se - TypeCraft Bookmarks Tool Act D TC I	o <b>∽</b> - Mozilla s <u>H</u> elp cditor	Firefox				ू प्रि	Nor:			е 🦳 Ма	* ♀ ■ slGram	▼ typec	raft.org		) 📰 Multilan
Interference Direct Norw	The interfined Closed Text Repository Provide N <td>) TC Editor - Mozilla F plications Places Sy</td> <td>irefox stem S 2 ? se - TypeCraft Bookmarks Tool Act D TC I</td> <td>o<b>∽</b> - Mozilla s <u>H</u>elp cditor</td> <td>Firefox</td> <td></td> <td>] Search</td> <td> <b>ж</b>от</td> <td></td> <td></td> <td><u>&amp;</u> Lars</td> <td>Hellan</td> <td>e 🦳 Ma my talk</td> <td>* ♀ ■ slGram</td> <td>▼ typec</td> <td>raft.org</td> <td></td> <td> ) 📰 Multilan</td>	) TC Editor - Mozilla F plications Places Sy	irefox stem S 2 ? se - TypeCraft Bookmarks Tool Act D TC I	o <b>∽</b> - Mozilla s <u>H</u> elp cditor	Firefox		] Search	<b>ж</b> от			<u>&amp;</u> Lars	Hellan	e 🦳 Ma my talk	* ♀ ■ slGram	▼ typec	raft.org		) 📰 Multilan
special page   we text alining templates prests   we text alining templates prests   search Phrase   search result (2 phrases found):	craft editor       special page         cew text raining templates by texts       Search Phrase         craft search harse search craft help       Language       Phrase       Translation       Contributor       Last changee         OS tags list loos tags list experial pages       Norwegian Bulgarian       Du har ingen rett til å holde meg ansvarlig Bokmål       You do not have the right to hold me responsible for the past.       Translation       Translation       Translation       Last changee         OS tags list loos tags list experial pages       Bulgarian       Hamau ripaeo ga MM gbp#Mull CMETKa 3a MHHADTO.       You do not have the right to hold me responsible for the past.       Translation       T	TC Editor - Mozilla F plications Places Sy	irefox stem S 2 ? se - TypeCraft Bookmarks Tool Act D TC I	o <b>∽</b> - Mozilla s <u>H</u> elp cditor	Firefox		) Search जंगल में मो जंगल	🛠 🦲 ा ।र नाचा किस   में	ने देखा? "। ्रमोर	Who saw the	<u>2</u> Lars peacock dance	Hellan in the fore	e Ma my talk ग?" किसने ा	।k २ व् ∎ alGram ि े र € ि my preference	▼ typec	raft.org		) 📰 Multilan
ew text aning templates y texts   craft search traft s	ew text aining templates y texts craft search hrase search craft search hrase search craft telp Do tags list ense tags list ELP Pages duisory Board uin Page duisory Board turrent Events ch mediawiki as search box       Search Phrase Language Phrase Du har ingen rett til å holde meg ansvarlig for fortiden. Du har ingen rett til å holde meg ansvarlig bokmål       You do not have the right to hold me responsible for the past. You do not have the right to hold me responsible for the past.       BIB-TC / Lars Hellan BIB-TC / Lars Hellan DU har ingen rett til å holde meg ansvarlig bokmål       You do not have the right to hold me responsible for the past.       BIB-TC / Lars Hellan BIB-TC / Lars Hellan DU har ingen rett til å holde meg ansvarlig bokmål       You do not have the right to hold me responsible for the past.       BIB-TC / Lars Hellan BIB-TC / Lars Hellan DU har ingen rett til å holde meg ansvarlig bokmål       You do not have the right to hold me responsible for the past.       BIB-TC / Lars Hellan DI har ingen DU har ingen rett til å holde meg ansvarlig for the past.       You do not have the right to hold me responsible for the past.       BIB-TC / Lars Hellan DU har ingen DU har ingen rett til å holde meg ansvarlig for the past.       You do not have the right to hold me responsible for the past.       BIB-TC / Lars Hellan DU har ingen DU har ing	TC Editor - Mozilla F olications Places Sy Search Phras Edit View History E Mano - n Imaga typecraft.org/tc2	irefox stem S ???? se - TypeCraft 200kmarks Iool Act D TC I wiki/Special:Typ	• <b>Mozilla</b> s <u>H</u> elp :ditor	Firefox		) Search जंगल में मो जंगल jaṅgala <i>forest</i> .	💥 🔄 T Iरनाचा किस अट.अट ने ांग.	ने देखा? "   मोर in   mõn io   pea	Who saw the	<u>ि</u> Lars peacock dance नाचा nāca ā dance	Hellan in the fore. ft k	e) — Ma my talk my talk वस्यने व वस्यने व वस्यने व वस्यने व	slGram	▼ typec	raft.org		) 📰 Multilan
Search result (2 phrases found):     arraft help   D5 tags list   Oss tags list   oss tags list   ense tags list   ELP Pages   pe IPA symbols   c Previous Page: 1 v of 1 Next > Number of results per page: 20 v Export in HTML Export in LaTeX Export in XML Prepare for wiki	search result (2 phrases found): search rase search rafe search rafe search rafe telp Do Language Phrase I Language Phrase Du har ingen rett til å holde meg ansvarlig for fortie nas Du har ingen rett til å holde meg ansvarlig Bokmål for fortiden. Do Stags list cos tags list tel Page Search result (2 phrases found): Search result (2 phrase found): Search result (2 phrases found): Search result (2 phrase	TC Editor - Mozilla F olications Places Sy Search Phras dit View History E Vaлo - п Ппад typecraft.org/tc2	irefox stem S V ? se - TypeCraft 300kmarks Tool Act D TC I wiki/Special:Typ Glossed Text I	Callar - Mozilla s Help Editor ECraft/Sear	Firefox		) Search जंगल में मो जंगल jaṅgala <i>forest</i> .	💥 🔄 T Iरनाचा किस अट.अट ने ांग.	ने देखा? "   मोर in   mõn io   pea	Who saw the	<u>ि</u> Lars peacock dance नाचा nāca ā dance	Hellan in the fore. ft k	e) — Ma my talk my talk वस्यने व वस्यने व वस्यने व वस्यने व	slGram	▼ typec	raft.org		) 📰 Multilan
search xt search xt search xrase search rank teight so hold me responsible bokmål of for fortiden. Du har ingen rett til å holde meg ansvarig bokmål of for fortiden. Du har ingen rett til å holde meg ansvarig bokmål of for fortiden. Du har ingen rett til å holde meg ansvarig bokmål of for fortiden. Bulgarian Hsmatu npaeo да ми държищ сметка за миналото. Previous Page 1 of 1 Next > Number of results per page: 20 v Export in HTML Export in LaTEX Export in XML Prepare for wiki	Search result (2 phrases found): Traft search traft search traft search Translation Contributor Last Contributor Last Bokmäl Du har ingen rett til å holde meg ansvarig Bokmäl Du har ingen rett til å holde meg ansvarig Bokmäl Du har ingen rett til å holde meg ansvarig Bokmäl Du har ingen rett til å holde meg ansvarig Bokmäl Du har ingen rett til å holde meg ansvarig Bokmäl Du har ingen rett til å holde meg ansvarig Bokmäl Du har ingen rett til å holde meg ansvarig Bokmäl Du har ingen rett til å holde meg ansvarig Bokmäl Du har ingen rett til å holde meg ansvarig Bokmäl Du har ingen rett til å holde meg ansvarig Bokmäl Du har ingen rett til å holde meg ansvarig Prepare for kellts per page: 20 v Export in HTML Export in LaTeX Export in XML Prepare for wikl se search ox	TC Editor - Mozilla R blications Places Sy Search Phras Edit View History E Wano - n Image Wano - n Image Marcha - n Imag	irefox stem S V ? se - TypeCraft 200kmarks Tool Act D TC I wiki/Special:Typ Glossed Text I special p	Craft/Sear Repository	Firefox		) Search जंगल में मो जंगल jaṅgala <i>forest</i> .	💥 🔄 T Iरनाचा किस अट.अट ने ांग.	ने देखा? "   मोर in   mõn io   pea	Who saw the	<u>ि</u> Lars peacock dance नाचा nāca ā dance	Hellan in the fore. ft k	e) — Ma my talk my talk वस्यने व वस्यने व वस्यने व वस्यने व	slGram	▼ typec	raft.org		) 📰 Multilan
Image: Language Phrase Translation Contributor Language   Image: Norwegian Du har ingen rett til åholde meg ansvarig for dor the past. Pour do not have the right to hold me responsible   Image: Language Du har ingen rett til åholde meg ansvarig for do not have the right to hold me responsible PBrB-C/ Lars Hellan   Image: Language Du har ingen rett til åholde meg ansvarig for do not have the right to hold me responsible PBrB-C/ Lars Hellan   Image: Language Du har ingen rett til åholde meg ansvarig for do not have the right to hold me responsible PBrB-C/ Lars Hellan   Image: Language Du har ingen rett til åholde meg ansvarig for do not have the right to hold me responsible PBrB-C/ Lars Hellan   Image: Language Norwegian Du har ingen rett til åholde meg ansvarig for do not have the right to hold me responsible PBrB-C/ Lars Hellan   Image: Language Previous Previous Page: Image: Im	Language       Phrase       Translation       Contributor       Last change         ast search brase search DS tags list closs tags list ense tags list ELP Pages pre IPA symbols       Du har ingen rett til åholde meg ansvarig Bokmål       You do not have the right to hold me responsible for the past.       BIB-TC / Lers Hellan       2015-04-111 cm         DS tags list closs tags list ense tags list ELP Pages pre IPA symbols       Bulgarian       Hamau npago da Mu dbpжulu CMETKa 3a MuHanoTo.       You do not have the right to hold me responsible for the past.       BIB-TC / Petya Denova       2015-04-111 cm         anse tags list ELP Pages pre IPA symbols       Keppert in HTML       Export in LaTeX       Export in XML       Pages I = 0 f 1       Next >       Vision of 1       Next >         Number of results per page:       Z0 = v       Vision of 1       Next >       Vision of 1       Next >       Vision of 1       Vision of 1       Next >       Vision of 1	TC Editor - Mozilla F dications Places Sy Search Phras dit View History E vano - n Imagu vano - n	irefox stem S V ? se - TypeCraft 200kmarks Tool Act D TC I wiki/Special:Typ Glossed Text I special p	Craft/Sear Repository	Firefox		) Search जंगल में मो जंगल jaṅgala <i>forest</i> .	💥 🔄 T Iरनाचा किस अट.अट ने ांग.	ने देखा? "   मोर in   mõn io   pea	Who saw the	<u>ि</u> Lars peacock dance नाचा nāca ā dance	Hellan in the fore. ft k	e) — Ma my talk my talk वस्यने व वस्यने व वस्यने व वस्यने व	slGram	▼ typec	raft.org		) 📰 Multilan
rase search     Language     Phrase     Iranslation     Contributor     changed       rath help     Norwegian     Du har ingen rett til å holde meg ansvalig Bokmål     You do not have the right to hold me responsible for the past.     BTB-TC / Lars Hellan     2015-04-17 10-00       DS tags list coss tags list ense tags list ELP Pages pe IPA symbols     Bulgarian     Hamau npabo да ми държиш сметка за миналото.     You do not have the right to hold me responsible for the past.     BTB-TC / Lers Hellan     2015-04-11 12-00       Vou do not have the right to hold me responsible pe IPA symbols     FtB-TC / Petya     2015-04-11 12-00     Osenova     2015-04-11 12-00       Ain Page     Vou do not have the right to hold me responsible pe IPA symbols     FtB-TC / Petya     2015-04-11 12-00       Ain Page     Vou do not have the right to hold me responsible for the past.     BTB-TC / Petya     2015-04-11 12-00       Ain Page     Vou do not have the right to hold me responsible for the past.     BtB-TC / Petya     2015-04-11 12-00       Visory Board urrent Events     Export in HTML     Export in LaTeX     Export in XML     Prepare for wiki	rase search     Language     Phrase     Iransation     Contributor     changed       ansae search     Norwegian     Du har ingen rett til å holde meg ansvarig Bokmål     You do not have the right to hold me responsible Bokmål     BTB-TC / Lars Hellan     2015-04-17 10: 0 cenova       DS tags list oss tags list ense tags list ELP Pages pe IPA symbols     Bulgarian     Hямаш право да ми държиш сметка за миналото.     You do not have the right to hold me responsible for the past.     BTB-TC / Petva Osenova     2015-04-11 12: 0 cenova       raft information     Kanged      Previous     Page     I v of 1     Next >       Number of results per page:     Z0 v        Number of results per page:     Z0 v       Export in HTML     Export in LaTeX     Export in XML           Prepare for wiki     Search            ox	TC Editor - Mozilla F dications Places Sy Search Phras dit View History E vano - n Image typecraft.org/c2	irefox stem S V ? se - TypeCraft 200kmarks Ioo Act D TC I wiki/Special:Typ Glossed Text I special p Searc	- Mozilla s Help ditor contr/sear ecraft/sear kepository age h Phr	Firefox Problem chPhrase/ ase	- <b>1</b> 0	] Search जंगल में मो जंगल jaṅgala <i>forest.</i> N	💥 🔄 T Iरनाचा किस अट.अट ने ांग.	ने देखा? "   मोर in   mõn io   pea	Who saw the	<u>ि</u> Lars peacock dance नाचा nāca ā dance	Hellan in the fore. ft k	e) — Ma my talk my talk वस्यने व वस्यने व वस्यने व वस्यने व	slGram	▼ typec	raft.org		) 📰 Multilan
rraft help       Bokmåi       for fortiden.       for the past       bitch / tals nemal       2015-04-11 12:0         DS tags list loss tags list mese tags list ELP Pages pe IPA symbols       Bulgarian       Hsmau npaso да ми държиш CMETKa за MUHADOTO.       You do not have the right to hold me responsible for the past.       BTB-TC / Petya Osenova       2015-04-1112:0         V ubber of results per page: 20 v        Previous       Page: 1 v of 1       Next >         Number of results per page: 20 v        Export in LaTeX       Export in XML         Prepare for wiki       Prepare for wiki	traft help DS tags list Bulgarian Нямаш право да ми държиш сметка за MuHaлото. Previous Page 1 v of 1 Next > Number of results per page: 20 v Export in HTML Export in LaTeX Export in XML Prepare for wiki Search ox	TC Editor - Mozilla F dications Places Sy Search Phras dit View History E vano - n Imag vano - n Imag	irefox stem S ? ? se - TypeCraft 200kmarks Iool Act D TC I wiki/Special:Typ Clossed Text I Special p Search Search re	A Mozilla s Help ditor ecraft/Sear Repository age h Phr esult (2 p	Firefox Problem chPhrase/ ase	- <b>1</b> 0	) Search अंगल में मो जंगल jangala <i>forest.</i> N	¥ा ार नाजा किस   मे   मन   मन	ने देखा? "   मोर in   mõn io   pea	Who saw the	<u>ि</u> Lars peacock dance नाचा nāca ā dance	Hellan in the fore: ft k k P	e Ma my taik st?" i isu në o oho a N i	= * ♀ ■ alGram ☆ ▼ € wy preference textual textu	▼ typec	raft.org watchlist r	my contr	- Multilan
DS tags list loss tags list loss tags list loss tags list lane tags lane tags land tags lane tags land tags land tags lane tags land tags land ta	DS tags list loss tags list ense tags list ELP Pages pre IPA symbols traft information ain Page dvisory Board durrent Events ch mediawiki ao so	TC Editor - Mozilla F olications Places Sy Search Phrase Edit View History E vano - n Imagu typecraft.org/tc2 typecraf	stem S ? ? se - TypeCraft 200kmarks Iool Act D TC I wiki/Special:Typ Clossed Text I Search ru Search ru	A Mozilla s Help ditor ecraft/Sear Repository age h Phr esult (2 p anguage	Firefox Problem chPhrase/ ase hrases 1	4 n Io ( found):	े Search अंगल में मो जंगल jangala forest. N	¥ा र नाथा किस े से मि मि मि मि मि मि मि मि मि मि मि मि मि	ने देखा? "'   मोर itin pea pea OST N	Who saw the	<u>े</u> Lars peacock dance । नामा năca a dance PTCP	Hellan in the fore. F N	e Ma my talk अर्थ भेक व भेषा भेषा भेषा भेषा भेषा भेषा भेषा भेषा	= * ♀ ■ alGram ☆ ▼ ♥ wy preference ter tokha a ter tokha a ter tokha a	▼ typeci	raft.org watchlist r	my contr	- Multilan
loss tags list ane tags list ELP Pages pe IPA symbols traft information ain Page tvisory Board urrent Events ch mediawkil io Search	loss tags list ense tags list ense tags list ELP Pages pe IPA symbols rraft information ain Page tylisory Board urrent Events ch mediawiki aio Search ox	TC Editor - Mozilla F dications Places Sy Search Phras dit View History E vano - п Imag vano - n Imag	irefox stem S ? ? ? se - TypeCraft gookmarks Tool Act D TC I wiki/Special:Typ Glossed Text I Special p Search re Search re 	- Mozilla s Help ditor ecraft/Sear ecraft/Sear h Phr esult (2 p anguage egian	Firefox Problem chPhrase/ ase bhrases 1 Du h	found): ar inger	े Search अंगल में मो जंगल jangala forest. N	¥ा र नाथा किस े से मि मि मि मि मि मि मि मि मि मि मि मि मि	ने देखा? "'   मोर itin pea pea OST N	Who saw the	<u>ि</u> Lars peacock dance नाचा गोदस ॥ dance PTCP	Hellan in the fore. F N	e Ma my talk अर्थ भेक व भेषा भेषा भेषा भेषा भेषा भेषा भेषा भेषा	= * ♀ ■ alGram ☆ ▼ ♥ wy preference ter tokha a ter tokha a ter tokha a	▼ typeci	raft.org watchlist r Contri	ny contr	Multilan
ELP Pages   pe IPA symbols	<pre>ELP Pages</pre>	TC Editor - Mozilla F Alications Places Sy Search Phras Edit View Higtory E vano - n Image typecraft.org/tc2 Typecraft.org/tc2 The Interlinear craft editor aw text aining templates y texts craft search arase search arase search craft help DS tags list	irefox stem S ? ? ? se - TypeCraft 200kmarks Iool Act D TC I wiki/Special:Typ Glossed Text I Search re Search re Search re Search re Search re Search re	- Mozilla s Help ditor ecraft/Sear ecraft/Sear h Phr esult (2 p anguage egian ål	Firefox Problen chPhrase/ ase ase hrases bhrases Du h for for Hawa	found): ar inger ortiden.	ु Search जंगल में मो जंगलुवाa <i>forest.</i> N P n rett til है	¥ ा र नापा दिस   म   म   म   म     म                 	मे देखा? " मोर मार (pea OST N	Who saw the ra cock	A Lars peaceck dance Then Then a dance PTCP	Hellan in the fore f P P The second s	e Ma my talk औ?" किसने के isa në ( isa në (	slGram slGram	v typect	raft.org watchlist r Contri BTB-TC / Lars BTB-TC / Lars	ny contr butor	Last
raft information   ain Page   dvisory Board   urrent Events   Search	raft information       Image: Search ox         So       Search	TC Editor - Mozilla F olications Places Sy Search Phras Edit View History E Pano - n Image typecraft.org/tc2 The Interlinear craft editor ew text aining templates y texts craft search txt search trase search craft help DS tags list loss tags list	irefox stem S ? ? ? se - TypeCraft 200kmarks Iool Act D TC I wiki/Special:Typ Glossed Text I Search re Search re Search re Search re Search re Search re	- Mozilla s Help ditor ecraft/Sear ecraft/Sear h Phr esult (2 p anguage egian ål	Firefox Problen chPhrase/ ase ase hrases bhrases Du h for for Hawa	found): ar inger ortiden.	ु Search जंगल में मो जंगलुवाa <i>forest.</i> N P n rett til है	¥ ा र नापा दिस   म   म   म   म     म                 	मे देखा? " मोर मार (pea OST N	Who saw the ra cock	A Lars peaceck dance Then Then a dance PTCP	Hellan in the fore f P P The second s	e Ma my talk औ?" किसने के isa në ( isa në (	slGram slGram	v typect	raft.org watchlist r Contri BTB-TC / Lars BTB-TC / Lars	ny contr butor	Last changed
ain Page   byisory Board   airrent Events   Prepare for wiki	Bain Page   dvisory Board   urrent Events	TC Editor - Mozilla F plications Places Sy Search Phras Edit View Higtory E parano - n Image typecraft.org/tc2 typecraft.o	irefox stem S ? ? ? se - TypeCraft 200kmarks Iool Act D TC I wiki/Special:Typ Glossed Text I Search re Search re Search re Search re Search re Search re	- Mozilla s Help ditor ecraft/Sear ecraft/Sear h Phr esult (2 p anguage egian ål	Firefox Problen chPhrase/ ase ase hrases bhrases Du h for for Hawa	found): ar inger ortiden.	ु Search जंगल में मो जंगलुवाa <i>forest.</i> N P n rett til है	¥ ा र नापा दिस   म   म   म   म     म                 	ने देखा? "। मोर हिस्स () () () () () () () () () () () () ()	Who saw the ra cock varlig	A Lars peacock dance states in dance prop prop	Hellan in the fore. F k y F T ave the i	e Ma my talk भर? केंद्रने में केंद्रने में केंद्रने में san në d san san san tanslatio right to h right to h	slGram slGram	v typect	raft.org watchlist r Contri BTB-TC / Lars BTB-TC / Lars	ny contr butor	Last changed
in mediawiki Search	dvisory Board urrent Events Go Search ox	TC Editor - Mozilla F olications Places Sy Search Phrase Edit View History E Pagno - n Image typecraft.org/tc2 Typecraft.org/tc2 The Interlinear craft editor ew text aning templates by texts craft search trase search trase search craft help OS tags list loss tags list ense tags list ELP Pages type IPA symbols	irefox stem S ? ? ? se - TypeCraft 200kmarks Ioo Act D TC ( wiki/Special:Typ Glossed Text ( Search re Search re Search re Search re Bulga Number	Aozilla     Mozilla     Leip ditor     craft/Sear      craft/Sear      control     co	Firefox Problem chPhrase/ a SC bhrases f bhrases f for fo Hsma MuHa s per pag	found): ar inger artinger anoto. ge: 20	ु Search जंगल में मो jorigala forest. N Pl n rett til å во да ми	¥ т к नापा रिस	ने देखा? " मो pea ost N neg ansv c cmetka < Pre	Who saw the ra cock varlig	A Lars peacock dance states in dance prop prop	Hellan in the fore. F k y F T ave the i	e Ma my talk भर? केंद्रने में केंद्रने में केंद्रने में san në d san san san tanslatio right to h right to h	slGram slGram	v typect	raft.org watchlist r Contri BTB-TC / Lars BTB-TC / Lars	ny contr butor	Last changed
ch mediawiki	urrent Events ch mediawiki Go Search ox	TC Editor - Mozilla F olications Places Sy Search Phras Edit View History E Pano - n Image typecraft.org/tc2 TC Editor typecraft.org/tc2 The Interlinear craft editor ew text aning templates y texts craft search traft search traft search traft search craft help DS tags list loss tags list ELP Pages pre IPA symbols craft information	irefox stem S ??? se - TypeCraft 200kmarks [00] Act D TC ( wiki/Special:Typ Clossed Text ( Search ro Search ro Search ro Norw Bulga Number Export	Anozilla     S Help      ditor     craft/Sear      tepository      age      h Phr      esult (2 p      anguage      egian      frian      of results	Firefox Problem chPhrase/ a SC bhrases f bhrases f for fo Hsma MuHa s per pag	found): ar inger artinger anoto. ge: 20	ु Search जंगल में मो jorigala forest. N Pl n rett til å во да ми	¥ т к नापा रिस	ने देखा? " मो pea ost N neg ansv c cmetka < Pre	Who saw the ra cock varlig	A Lars peacock dance states in dance prop prop	Hellan in the fore. F k y F T ave the i	e Ma my talk भर? केंद्रने में केंद्रने में केंद्रने में san në d san san san tanslatio right to h right to h	slGram slGram	v typect	raft.org watchlist r Contri BTB-TC / Lars BTB-TC / Lars	ny contr butor	Last changed
so Search	Go Search	C Editor - Mozilla F olications Places Sy Search Phras Edit View History E avano - n Image typecraft.org/tc2 Typecraft.org/tc2 Typecraft.org/tc2 The Interlinear craft editor ew text aning templates ty texts craft search ext search craft help OS tags list loss tags l	irefox stem S ??? se - TypeCraft 200kmarks [00] Act D TC ( wiki/Special:Typ Clossed Text ( Search ro Search ro Search ro Norw Bulga Number Export	Anozilla     S Help      ditor     craft/Sear      tepository      age      h Phr      esult (2 p      anguage      egian      frian      of results	Firefox Problem chPhrase/ a SC bhrases f bhrases f for fo Hsma MuHa s per pag	found): ar inger artinger anoto. ge: 20	ु Search जंगल में मो jorigala forest. N Pl n rett til å во да ми	¥ т к नापा रिस	ने देखा? " मो pea ost N neg ansv c cmetka < Pre	Who saw the ra cock varlig	A Lars peacock dance states in dance prop prop	Hellan in the fore. F k y F T ave the i	e Ma my talk भर? केंद्रने में केंद्रने में केंद्रने में san në d san san san tanslatio right to h right to h	slGram slGram	v typect	raft.org watchlist r Contri BTB-TC / Lars BTB-TC / Lars	ny contr butor	Last changed
	xox	TC Editor - Mozilla F plications Places Sy Search Phrase Edit View History E Edit View History E typecraft.org/tc2 Typecraft.org/tc2 Typecraft.org/tc2 The Interlinear craft editor tew text ranining templates ty texts craft search trase search craft search trase search craft search craft search craft search sease list tense tags l	irefox stem S ??? se - TypeCraft 200kmarks [00] Act D TC ( wiki/Special:Typ Clossed Text ( Search ro Search ro Search ro Norw Bulga Number Export	Anozilla     S Help      ditor     craft/Sear      tepository      age      h Phr      esult (2 p      anguage      egian      frian      of results	Firefox Problem chPhrase/ a SC bhrases f bhrases f for fo Hsma MuHa s per pag	found): ar inger artinger anoto. ge: 20	ु Search जंगल में मो jorigala forest. N Pl n rett til å во да ми	¥ т к नापा रिस	ने देखा? " मो pea ost N neg ansv c cmetka < Pre	Who saw the ra cock varlig	A Lars peacock dance states in dance prop prop	Hellan in the fore. F k y F T ave the i	e Ma my talk भर? केंद्रने में केंद्रने में केंद्रने में san në d san san san tanslatio right to h right to h	slGram slGram	v typect	raft.org watchlist r Contri BTB-TC / Lars BTB-TC / Lars	ny contr butor	Last changed
	xoo	<ul> <li>C Editor - Mozilla F</li> <li>plications Places Sy</li> <li>Search Phrase</li> <li>Edit View History E</li> <li>typecraft.org/cc2</li> <li>typecra</li></ul>	irefox stem S ??? se - TypeCraft 200kmarks [00] Act D TC ( wiki/Special:Typ Clossed Text ( Search ro Search ro Search ro Norw Bulga Number Export	Anozilla     S Help      ditor     craft/Sear      tepository      age      h Phr      esult (2 p      anguage      egian      frian      of results	Firefox Problem chPhrase/ a SC bhrases f bhrases f for fo Hsma MuHa s per pag	found): ar inger artinger anoto. ge: 20	ु Search जंगल में मो jorigala forest. N Pl n rett til å во да ми	¥ т к नापा रिस	ने देखा? " मो pea ost N neg ansv c cmetka < Pre	Who saw the ra cock varlig	A Lars peacock dance states in dance prop prop	Hellan in the fore. F k y F T ave the i	e Ma my talk भर? केंद्रने में केंद्रने में केंद्रने में san në d san san san tanslatio right to h right to h	slGram slGram	v typect	raft.org watchlist r Contri BTB-TC / Lars BTB-TC / Lars	ny contr butor	Last changed
		TC Editor - Mozilla F  plications Places Sy  C Search Phras Edit View History E  range typecraft.org/tc2  t	irefox stem S ??? se - TypeCraft 200kmarks [00] Act D TC ( wiki/Special:Typ Clossed Text ( Search ro Search ro Search ro Norw Bulga Number Export	Anozilla     S Help      ditor     craft/Sear      tepository      age      h Phr      esult (2 p      anguage      egian      frian      of results	Firefox Problem chPhrase/ a SC bhrases f bhrases f for fo Hyma MuHa s per pag	found): ar inger artinger anoto. ge: 20	ु Search जंगल में मो jorigala forest. N Pl n rett til å во да ми	¥ т к नापा रिस	ने देखा? " मो pea ost N neg ansv c cmetka < Pre	Who saw the ra cock varlig	A Lars peacock dance states in dance prop prop	Hellan in the fore. F k y F T ave the i	e Ma my talk भर? केंद्रने में केंद्रने में केंद्रने में san në d san san san tanslatio right to h right to h	slGram slGram	v typect	raft.org watchlist r Contri BTB-TC / Lars BTB-TC / Lars	ny contr butor	Last changed
		TC Editor - Mozilla F olications Places Sy Search Phrase Edit View History E Pano - n Image typecraft.org/tc2 TC Editor end typecraft.org/tc2 TC Editor end typecraf	irefox stem S ??? se - TypeCraft 200kmarks [00] Act D TC ( wiki/Special:Typ Clossed Text ( Search ro Search ro Search ro Norw Bulga Number Export	Anozilla     S Help      ditor     craft/Sear      tepository      age      h Phr      esult (2 p      anguage      egian      frian      of results	Firefox Problem chPhrase/ a SC bhrases f bhrases f for fo Hyma MuHa s per pag	found): ar inger artinger anoto. ge: 20	ु Search जंगल में मो jorigala forest. N Pl n rett til å во да ми	¥ т к नापा रिस	ने देखा? " मो pea ost N neg ansv c cmetka < Pre	Who saw the ra cock varlig	A Lars peacock dance states in dance prop prop	Hellan in the fore. F k y F T ave the i	e Ma my talk भर? केंद्रने में केंद्रने में केंद्रने में san në d san san san tanslatio right to h right to h	slGram slGram	v typect	raft.org watchlist r Contri BTB-TC / Lars BTB-TC / Lars	ny contr butor	Last changed
		<ul> <li>TC Editor - Mozilla F</li> <li>plications Places Sy</li> <li>Search Phras</li> <li>Edit View History E</li> <li>typecraft.org/tc2</li> <li>typecraft.org/</li></ul>	irefox stem S ??? se - TypeCraft 200kmarks [00] Act D TC ( wiki/Special:Typ Clossed Text ( Search ro Search ro Search ro Norw Bulga Number Export	Anozilla     S Help      ditor     craft/Sear      tepository      age      h Phr      esult (2 p      anguage      egian      frian      of results	Firefox Problem chPhrase/ a SC bhrases f bhrases f for fo Hyma MuHa s per pag	found): ar inger artinger anoto. ge: 20	ु Search जंगल में मो jorigala forest. N Pl n rett til å во да ми	¥ т к नापा रिस	ने देखा? " मो pea ost N neg ansv c cmetka < Pre	Who saw the ra cock varlig	A Lars peacock dance states in dance prop prop	Hellan in the fore. F k y F T ave the i	e Ma my talk भर? केंद्रने में केंद्रने में केंद्रने में san në d san san san tanslatio right to h right to h	slGram slGram	v typect	raft.org watchlist r Contri BTB-TC / Lars BTB-TC / Lars	ny contr butor	Last changed

### 3. Preliminary annotation of MWEs with catena in lexicons.

As a model we use the way in which verbal MWEs are represented in Norsource, exemplified below with the lexical entry for *alliere* ('ally') for the frame where it selects a reflexive and a PP headed by the preposition *med* (corresponding to English 'ally oneself with'); cf. (3) below. Like all verb entries, this one consists of an orthographic form (conventionally selected to be like the infinitive form), under STEM, a conjugation class under INFLECTION indicator ('nonfstr-te' representing 'non-finally-stressed infinitival form, of the type selecting 'te' as past tense form), and a semantic predicate placeholder (introduced by the path 'SYNSEM.LKEYS.KEYREL.PRED'). In addition the entry has the line 'SYNSEM.LOCAL.KEY-SPEC med-assoc' whereby 'med' is indicated as a selected preposition. (In the source format notation used, ':=' means 'is a subtype of' and '&' means 'unified with'.)

(3)
alliere\_tr-obl-refl := v-trObl-obRefl\_oblN &
 [ INFLECTION nonfstr-te,
 STEM < "alliere" >,
 SYNSEM.LKEYS.KEYREL.PRED "\_alliere\_v-trObl\_rel",
 SYNSEM.LOCAL.KEY-SPEC med-assoc ].

The specification in the last line in the lexical entry is integrated into the general type system of the grammar through the type definition below, where the path 'LOCAL.KEY-SPEC \_' is defined as having the same value as 'LOCAL.CAT.VAL.ICOMPS <[LOCAL.CAT.HEAD [KEYS.KEY \_ ]]>' (indicated by '#1' in both underscore positions), where the longer path specifies restrictions on a selected PP for the verb frame in question:

(4)

trans-obl-synsem-sup := arg1-subj-synsem & arg2-comps-synsem & [LOCAL.CAT.HEAD non-copula, LOCAL.CONT.HOOK hook & [VARG #vrg], LOCAL.CAT trans-with-licomps-pp-cat, LOCAL.CAT.VAL.COMPS <[LOCAL.CAT.HEAD nominal ]>, LOCAL.KEY-SPEC #1, LOCAL.CAT.VAL.ICOMPS <[LOCAL.CAT.HEAD prep-or-adv & [SELECTED +, KEYS.KEY #1 & index-sita]]>, LOCAL.CAT.VAL.ICOMPS <[LOCAL.CAT.HEAD [KEYS.KEY #1 ]]>, LOCAL.CAT.QVAL.OBL1.LOCAL.CONT.HOOK.LTOP #4 & handle, LOCAL.CAT.VAL.ICOMPS <[LOCAL.CAT.HEAD [SELECTED +]]>, LOCAL.CAT.VAL.ICOMPS <[LOCAL.CAT.HEAD [SELECTED +]]>, LOCAL.CAT.VAL.ICOMPS <[LOCAL.CAT.HEAD [SELECTED +]]>, LOCAL.CAT.VAL.ICOMPS <[LOCAL.CAT.HEAD [SELECTED +]]>,

Although such general definitions have to be given for all relevant verb frame types (which, supertypes included, may reach about 150 in a complete inventory), the strategy is straightforward to implement in any grammar with the feature structure format in question (reflecting the 'HPSG Grammar Matrix' design<sup>8</sup>), such as Norsource and BURGER.

4. Preliminary ideas on data exchange strategy between CLaRK<sup>9</sup> system and TypeCraft: Since BulTreeBank has been created in the CLaRK system, a two-direction conversion template will be created in order to manage the data between the two systems in XML.

<sup>&</sup>lt;sup>8</sup> See Bender et al. 2010.

<sup>&</sup>lt;sup>9</sup> See Simov et al 2001.

For the incremental addition of new data, the template will be used for both purposes: loading the data into TypeCraft, encoding it (with valencies and catena), and bridging it back to the CLaRK system format.

In this way, the communication between the two systems will be ensured for the purposes of control and data syncronization.

## **Future Work**

We plan to:

- 1. extend the parallel MWE corpus further with data from parallel corpora;
- 2. enrich the lexicons in the grammars with MWEs;
- 3. test the format of catena-based description on more data and more MWE types;
- 4. implement the data exchange strategy mentioned under 'Results', point 4;
- 5. relate the parallel MWE annotation to the 3D system of verb action representation,  $ImagAct^{10}$ ;

6. with a view to Machine Translation as a field of application of the encoding of MWEs in corpora and grammars, explore ways of reflecting MWEs in the semantic representation format Minimal Recursion Semantics ('MRS', cf. Copestake et al. 2005), which constitutes part of the 'matrix' architecture.

Joint publications are envisaged reflecting some of these planned activities.

A precise time plan for these activities has not been made yet; most of the cooperation can be done remotely between Sofia and Trondheim, but we envisage applying for a further STM to enhance aspects of the work calling for more direct cooperation.

# References

- Beermann, D and L. Hellan. 2004. A treatment of directionals in two implemented HPSG grammars. In Stefan Müller (ed) Proceedings of the HPSG04 Conference, Katholieke Universiteit Leuven. CSLI Publications /<u>http://csli-publications.stanford.edu/</u>
- Beermann, D. and Mihaylov, P. (2013). Collaborative databasing and Resource sharing for Linguists. *Languages Resources and Evaluation*. Springer, 1-23. Bender, E. M., Drellishak, S., Fokkens, A., Poulson, L. and Saleem, S. (2010). Grammar

- Customization. In *Research on Language & Computation*, Volume 8, Number 1, 23-72. Copestake, A. (2002). *Implementing Typed Feature Structure Grammars*. CSLI Publications. Copestake, A., Flickinger, D., Pollard, C., and Sag, I. (2005). Minimal Recursion Semantics. In *Journal of Language and Computation*. 281-332.
- Hellan, L. (2008) From Grammar-Independent Construction Enumeration to Lexical Types in Computational Grammars. Paper presented at COLING, Workshop on Grammar Engineering Across Frameworks (GEAF) Manchester, August 2008 (http://www.aclweb.org/anthologynew/W/W08/#1700).
- Hellan, L., D. Beermann, T. Bruland, M.E.K. Dakubu, and M. Marimon (2014) MultiVal: Towards
- Hellan, L., D. Beermann, T. Bruland, M.E.K. Dakubu, and M. Marimon (2014) MultiVal: Towards a multilingual valence lexicon. LREC 2014.
  Osenova, Petya Osenova (2011). Localizing a Core HPSG-based Grammar for Bulgarian. In: Hanna Hedeland, Thomas Schmidt, Kai Worner (eds.) *Multilingual Resources and Multilingual Applications, Proceedings of GSCL 2011*, ISSN 0176-599X, Hamburg, pp. 175-180.
  Pollard, C. and Sag, I. (1994). *Head-Driven Phrase Structure Grammar*. Chicago University Press. Simov, K., Osenova, P., Simov, A. and Kouylekov, M. (2005): Design and Implementation of the Bulgarian HPSG-based Treebank. In Erhard Hinrichs and Kiril Simov, editors, *Journal of Research on Language and Computation*, Special Issue, Springer, 2005, pp. 495-522.
  Simov, K and Osenova, P. (2014): Formalizing MultiWords as Catenae in a Treebank and in a Lexicon. In: Verena Henrich. Erhard Hinrichs. Daniël de Kok. Petya Osenova. Adam

- Lexicon. In: Verena Henrich, Érhard Hinrichs, Daniël de Kok, Petya Osenova, Adam Przepiórkowski (eds.) Proceedings of the Thirteenth International Workshop on Treebanks and Linguistic Theories (TLT13), December 12-13, 2014, Tübingen, Germany, pp. 198-207. Simov, K., Peev, Z., Kouylekov, M., Simov, A., Dimitrov, M., Kiryakov, A (2001). CLaRK - an
- XML-based System for Corpora Development. In: Proc. of the Corpus Linguistics 2001 Conference, pages: 558-560.

<sup>&</sup>lt;sup>10</sup> See <u>http://www.imagact.it</u>