

3rd International Conference on Meaning and Knowledge Representation

Universidad Politécnica de Valencia

2-4 July 2014,
Valencia (Spain)

SPONSORED BY:



UNIVERSITAT
POLITÈCNICA
DE VALÈNCIA



This conference has been co-financed by Generalitat Valenciana - Conselleria de Educació, Cultura y Deporte
(AORG/2014/071)

Table of Contents

CONFERENCE PROGRAM	p. 4
PLENARY SPEAKERS (Biodata).....	p. 8
ROUND-TABLE DISCUSSANTS (Biodata).....	p. 12
ABSTRACTS	
• Wednesday, 2 July 2014.....	p. 15
• Thursday, 3 July 2014.....	p. 22
• Friday, 4 July 2014.....	p. 31

Conference Program

**3rd International Conference on Meaning and Knowledge Representation
(2-4 July, 2014)**

-
Conference Program

Wednesday, 2 July 2014	
9:00	Registration
9:30	Opening Act
10:00	Plenary Lecture: Constructions and the lexicon in Functional Discourse Grammar <i>Lachlan Mackenzie</i>
11:00	Coffee Break
11:30	Argument realisation of the concepts TRANSFER, LET /ALLOW and PERMISSION in Modern Irish: The interaction of causation, event chaining and syntactic variation <i>Brian Nolan</i>
12:00	Causative 'lassen' constructions in German: Syntax, argument structure, meaning variants and the impact of cultural knowledge in disambiguation <i>Elke Diedrichsen</i>
12:30	The integration of transitive predicates into intransitive constructions: The case of the intransitive motion construction <i>Aneider Iza Erviti</i>
13:00	Towards the meaning and realization of Māori neuter verbs <i>Aoife Finn</i>
14:00	Lunch
15:30	Plenary Lecture: Cognitive modeling, motivation and explanation across descriptive levels in the Lexical Constructional Model <i>Francisco Ruiz de Mendoza Ibáñez</i>
16:30	A systematic approach to labelling external constraining factors in the Lexical Constructional Model <i>M. Sandra Peña Cervel</i>
17:00	Coffee Break
17:30	Formation of a substantive core of the polysemous verb of relations <i>Svetlana Kiseleva</i> <i>Nelly Trofimova</i>
18:00	Can conceptual integration optimise the activation of embedded scripts? <i>Alexey Yavetskiy</i>
18:30	Round-Table Discussion: What can theoretical linguistics do for NLP research? <i>Elke Diedrichsen, Brian Nolan,</i> <i>Carlos Periñán-Pascual, Francisco Rangel Pardo</i>

Thursday, 3 July 2014	
9:00	Plenary Lecture: Meaning and knowledge representation in anaphora resolution <i>Ruslan Mitkov</i>
10:00	Interpretation of co-referential chains in Czech <i>Alena Poncarová</i>
10:30	Automatic coreference resolution with FunGramKB <i>María José Ruiz Frutos</i>
11:00	Coffee Break
11:30	The contribution of FunGramKB to study semantic memory loss <i>María Beatriz Pérez Cabello de Alba</i> <i>Eva Samaniego Fernández</i>
12:00	An analysis of the prototypical senses of prepositions in, on and at to be implemented in an NLP system <i>Diana Hernández Pastor</i>
12:30	Building the rules for referential phrases within ARTEMIS <i>Francisco Cortés Rodríguez</i>
13:00	ARTEMIS: Designing the rules for do insertion in English <i>Ana Díaz Galán</i> <i>María del Carmen Fumero Pérez</i>
14:00	Lunch
15:30	Plenary Lecture: Methodological underpinnings for the construction of a criminal law subontology in FunGramKB <i>Ángel Felices Lago</i>
16:30	The integration of the concept +CRIME_00 in FunGramKB and the conceptualization or hierarchization problems involved. <i>Ángela Alameda Hernández</i> <i>María Enriqueta Cortés de los Ríos</i> <i>Ángel Felices Lago</i>
17:00	Coffee Break
17:30	The population of lexical units corresponding to Globalcrimeterm concepts in the FunGramKB's lexico-grammatical module <i>María Diana Fernández Lloret</i> <i>María Ángeles Gómez Castejón</i>
18:00	Procedural processes in Criminal Law: conceptualization of expert knowledge in FunGramKB Ontology <i>María de Gracia Carrión Delgado</i> <i>Alba Luzondo Oyón</i>
19:00	SOCIAL EVENT: Guided Tour
21:00	COCKTAIL: Muvim Restaurant (10 Quevedo Street)

Friday, 4 July 2014	
9:00	A polysemy account of Turkish spatial noun 'üst' in dative case marker <i>Aysun Balkan</i>
9:30	A hypothesis about origin of meta-symbols and superordinate categorization <i>Ciro Antunes de Medeiros</i>
10:00	Code-mixing and social identity: A visual approach <i>Jayshree Chakraborty</i>
10:30	An iconic and a systematic semantic feature of 'irregular' forms in English <i>Elena Even-Simkin</i>
11:00	Coffee Break
11:30	Interface of language and culture: the Mandaya phenomenon <i>Leorisyl D. Siarot</i>
12:00	Approaching illocutionary constructions within a computational environment <i>Alba Luzondo Oyón</i> <i>Rocío Jiménez Briones</i>
12:30	Finding a way through the population in FunGramKB with a Latent Semantic Analysis-based tool <i>María Beatriz Pérez Cabello de Alba</i> <i>Ismael Iván Teomiro García</i>
13:00	From Terminography to Ontology: The representation of legal concepts in FunGramKB <i>Gianluca Pontrandolfo</i> <i>Pedro Ureña Gómez-Moreno</i>
13:30	Closing Act

Plenary Speakers (Biodata)

Lachlan Mackenzie (VU University Amsterdam, Netherlands)

Lecture: Constructions and the lexicon in Functional Discourse Grammar

Date: Wednesday, 2 July 2014 (10:00)

Born in Scotland in 1950, Dr. J. Lachlan Mackenzie graduated in French and German Language and Literature and then went on to complete a PhD in General Linguistics in 1978 at Edinburgh. From 1977 to 1987 he was a Lecturer at VU University Amsterdam, where he later became a Full Professor in English Language and the Faculty's Director of Research; since 2007 he has been a part-time Professor of Functional Linguistics there. He is also a researcher at CELGA-ILTEC, a linguistics institute in Lisbon and Coimbra, Portugal. He is Research Manager of SCIMITAR, an international network for linguistics and discourse analysis based in Santiago de Compostela. In the first fifteen years of his presence in Amsterdam, he worked closely with Simon C. Dik on Functional Grammar. In more recent years he has collaborated with Kees Hengeveld on the elaboration of Functional Discourse Grammar, publishing Functional Discourse Grammar (OUP) and various articles with him; in addition, he has been active as an editor of the international journal of functional linguistics, *Functions of Language*. He has applied his linguistic knowledge in textbooks for students on practical grammar and academic writing, including *Compare and Contrast*, a grammar of English for Spanish speakers co-written with Elena Martínez Caro. His most recent book is *Casebook in Functional Discourse Grammar* (John Benjamins), co-edited with Hella Olbertz. See www.lachlanmackenzie.info

Francisco Ruiz de Mendoza Ibáñez (Universidad de La Rioja, Spain)

Lecture: Cognitive modeling, motivation and explanation across descriptive levels in the Lexical Constructional Model

Date: Wednesday, 2 July 2014 (15:30)

Dr. Francisco Ruiz de Mendoza Ibáñez is full professor on Linguistics at the University of La Rioja. He works in cognitive linguistics, inferential pragmatics, and functional grammar. He has published several books and well over 100 articles and book chapters. He has been invited -on more than fifty occasions- to present his research in universities and other research institutions in Madrid, Pisa, Galati, Timisoara, Hamburg, Copenhagen, Lodz, Las Palmas, Provo, Beijing, Lima, Bogotá, among other places. He has also been an invited speaker in national and international conferences, among them the International Conference on Cognitive Linguistics (Pavia, 2003), International Contrastive Linguistics Conference (Santiago, 2005) and the 10th International Cognitive Linguistics Conference (Krakow, 2007), III Conference on Metaphor in Language and Thought (Fortaleza, Ceará, Brazil, 2008), Languages and International Business: First Forum across Disciplines (Beijing, 2010), Cognitive Perspectives on Contrastive Grammar (Bielsko-Biala, Poland, 2011), First International Conference on Meaning and Knowledge Representation (Madrid, 2012). He serves on the editorial and scientific boards of a number of journals such as *Miscelánea*, *Estudios de Filología Moderna*, *Jezikoslovlje*, *Estudios Ingleses de la Universidad Complutense*, *Revista Alicantina de Estudios Ingleses*, *Revista Española de Lingüística Aplicada*, *IIL-International Journal of Applied Linguistics*, *International Journal of Cognitive Linguistics*, *Revue Romane*, and *Cognitive Linguistics*. He was the editor of the *Revista Española de Lingüística Aplicada* from 1998 to 2005 and is currently the editor of the *Review of Cognitive*

Linguistics, and co-editor of the series Applications of Cognitive Linguistics (Mouton de Gruyter). Since January 2012 he is co-editor of the Bibliography of Metaphor and Metonymy (Metbib), published by John Benjamins. He was president of the Spanish Association of Applied Linguistics from 2005 to 2011. He has been the head organizer of two major international conferences, the 8th International Cognitive Linguistics Conference (University of La Rioja, 2003) and the 44th International Meeting of the Societas Linguistica Europaea (University of La Rioja, 2011). He has also organized a number of other smaller conferences, such as the two Conferences of the Research Center on the Applications of Language, CRAL 2009, 2013, and several workshops and seminars.

Ruslan Mitkov (University of Wolverhampton, United Kingdom)

Lecture: Meaning and knowledge representation in anaphora resolution

Date: Thursday, 3 July 2014 (9:00)

Dr. Ruslan Mitkov's extensively cited research includes 200 publications on various topics of Natural Language Processing (NLP). His research output was highlighted as internationally leading in the last UK Research Assessment Exercise. Prof. Mitkov is best known for his seminal contributions to anaphora resolution and automatic generation of multiple-choice tests but his research has been influential in other areas too including centering, translation memory, evaluation, term extraction, cognates and false friends, and translation universals. Prof. Mitkov is author of the authoritative monograph Anaphora resolution (Longman) and sole Editor of one of the most successful handbooks published by Oxford University Press: The Oxford Handbook of Computational Linguistics. Current prestigious projects include his role as Executive Editor of the Journal of Natural Language Engineering (Cambridge University Press) and Editor-in-Chief of the NLP book series of John Benjamins publishers. Prof. Mitkov has been invited as a keynote speaker at a number of international conferences and has acted as Programme Chair of major conferences on NLP, Machine Translation, Translation Technology and Anaphora Resolution. He is regularly asked to review for prominent international funding bodies and organisations, for leading journals, publishers and conferences, to serve as a member of Programme Committees or Editorial Boards and to act as a referee for applications for Professorships in North America and Europe. Dr. Mitkov has considerable external funding to his credit and is currently managing several large projects, funded by UK research councils, by the EC as well as by companies and users from the UK and USA. Ruslan Mitkov received his university degree from the Humboldt University in Berlin, his PhD from the Technical University in Dresden and has worked as a Research Professor at the Institute of Mathematics, Bulgarian Academy of Sciences, Sofia. Mitkov is Professor of Computational Linguistics and Language Engineering, and Director of the Research Institute in Information and Language Processing at the University of Wolverhampton which he joined in 1995 and where he established the highly successful Research Group in Computational Linguistics. In recognition of his outstanding professional/research achievements, Prof. Mitkov was awarded the title of Doctor Honoris Causa at Plovdiv University in 2011.

Ángel Felices Lago (Universidad de Granada, Spain)

Lecture: Methodological underpinnings for the construction of a criminal law subontology in FunGramKB

Date: Thursday, 3 July 2014 (15:30)

Dr. Ángel Felices Lago began his university career at the University of Granada in 1984 and is currently Full Professor at the Department of English and German Philology. He teaches English and Spanish for specific purposes for undergraduate and postgraduate students at the Faculty of Business Studies and the Center for Modern Languages. He was Associate Dean for International Relations (1993-2001) and head of the Center for Modern Languages at the University of Granada (2001-2008). His main areas of research interest go from lexicology, discourse analysis and axiological linguistics to LSP teaching and knowledge engineering. He has co-authored or co-edited various scholar and pedagogical books and has also published over 80 scholarly articles and reviews in specialized national and international journals and volumes. He has served as invited reviewer or member on the editorial and scientific boards of a number of journals such as *Mediterranean Studies*, *Ibérica*, *Materiales para la enseñanza multicultural*, *Ideas*, *Cizí Jazyky*, *Cuadernos AISPI*, *Review of Cognitive Linguistics*, *Onomázein*, or *Odisea*. He has also led or taken part in various international academic projects funded by the Spanish Ministry of Science and Education and the European Union (Tempus, Leonardo, Erasmus, Erasmus-Mundus), and is currently leading a R&D project funded by the Spanish Ministry of Economy and Competitiveness to create a terminological ontology based on deep semantics, with the cooperation of 7 universities and international legal institutions. Finally, he has given lectures, special courses or conducted seminars in institutions from all over the world (Bratislava, Katowice, Madrid, Minsk, Moscow, Nijmegen, Prague, Salt Lake City, Santander (Menéndez Pelayo U.), Sao Paulo, Swansea, Trieste or Vienna, among others).

Round-Table Discussants (Biodata)

Elke Diedrichsen (Institute of Technology Blanchardstown, Ireland)

Dr. Elke Diedrichsen is a German linguist based in Dublin. She has worked as a researcher and deputy professor in several universities across Germany. Until recently, she worked as a Project manager in an international team of linguists for development and quality assurance with Google speech products (NLP, ASR, TTS) in Google's European Headquarters in Dublin, Ireland. She is a Member of the Computational and Functional Linguistics Research Group at the Institute of Technology Blanchardstown (ITB), Dublin. Elke Diedrichsen has widely published about functional linguistics, Role and Reference Grammar, constructions as grammatical objects, NLP, pragmatics and the semiotic, cultural and interactional potential of memes. Her recent publications include a volume co-edited with Brian Nolan (2013), entitled *Linking constructions into functional linguistics* – The role of constructions in grammars. She is also the author of the paper “A Role and Reference Grammar parser for German” which appeared in *Language processing and grammars: The role of functionally oriented computational models*, edited by Brian Nolan and Carlos Periñán-Pascual (2014). The software is able to classify German sentences and parse them into eight German sentence constructions. It achieves a lexical feature classification and displays the lexical features of the components (number, gender, case etc., also: human vs. inanimate). Elke Diedrichsen is currently co-editing a volume on argument structure in functional grammar with Brian Nolan and Gudrun Rawoens on *Causation, Permission and transfer – Argument realisation in GET, GIVE, PUT and TAKE verbs* to appear with Benjamins.

Brian Nolan (Institute of Technology Blanchardstown, Ireland)

Dr. Brian Nolan is Head of Department of Informatics and Creative Digital Media at the Institute of Technology Blanchardstown, Dublin in Ireland. The research interests of Dr. Nolan include computational approaches to language processing, linguistic theory at the morphosyntactic-semantic interface, argument structure, valence, event structure, and the architecture of the lexicon. His linguistic work has been in the functional linguistic model of Role and Reference Grammar (RRG) and he has published extensively internationally. Recent work includes the development of a rule-based Arabic to English machine translation engine with the RRG linguistic model supporting an interlingua bridge and the investigation of linguistic models to underpin Irish Sign Language avatars. Dr. Brian Nolan is the author of a book, published in 2012 by Equinox UK, on the linguistic structure of Irish in an RRG account entitled: ‘The structure of Irish: A functional account’. He co-edited two books of collected papers from international authors on ‘Linking Constructions into functional linguistics – The role of constructions in RRG grammar’ and ‘Language processing and grammars: The role of functionally oriented computational models’, published within the John Benjamins ‘Studies in Language’ Series. Dr. Nolan has 40 years experience nationally and internationally within the computer industry in a variety of senior roles and is also a widely published professional linguist.

Carlos Perriñán-Pascual (Universidad Politécnica de Valencia, Spain)

Dr. Carlos Perriñán-Pascual studied English Language and Literature at Universidad de Valencia (Spain) and received his Ph.D. degree in English Philology at UNED (Spain) in 1999. Since his doctoral dissertation on the resolution of word sense disambiguation in machine translation, his main research interests have been focused on knowledge engineering, natural language understanding and computational linguistics. As a result, he has been the director and founder of the FunGramKB project since 2004, whose main goal is to develop a lexico-conceptual knowledge base to be implemented in NLP systems requiring language comprehension. After having designed the knowledge base, he is now developing some tools for the FunGramKB Suite, such as the terminology extractor, the conceptual logical structure builder, and the inference engine. His scientific production includes about 50 peer-reviewed publications in the fields of linguistics, natural language processing and artificial intelligence. He is currently an Associate Professor in the Applied Linguistics Department at Universidad Politécnica de Valencia (Spain).

Francisco M. Rangel Pardo (Autoritas Consulting S.A., Spain)

Francisco Rangel is CTO at Autoritas Consulting and main scientist of the Social Business Intelligence tool Cosmos. Cosmos allows organizations to include social knowledge into their intelligence cycle by retrieving and processing millions of social media conversations in real time. Coordinator of Active Listening and Analysis areas at Socialancer, one of the main blogs about social media in Spain and Latin America. MAVIR 2007 Award for the best research work in Natural Language Processing & Information Retrieval. Dell 2009 Award Finalist in Technological Excellence (Fivasa Group/ SIC+). Mobip 2010 Award to the Most Innovative Idea (Corex SIC+). Co-organiser of the PAN task on Author Profiling in 2013 & 2014 at CLEF, whose aim has been identifying age and gender of authors of anonymous social media texts. On May, 17th Francisco Rangel has been interview at the "Informe Semanal" of TVE-1 where he described Autoritas's natural language processing activities on social media texts.

Abstracts

**Argument realisation of the concepts TRANSFER, LET/ALLOW and PERMISSION in
Modern Irish: The interaction of causation, event chaining and syntactic variation**

Brian Nolan

Computational and Functional Linguistics Research Group
Institute of Technology Blanchardstown, Dublin, IRELAND

This paper explores the encoding of the concepts of *transfer*, *let/allow* and *permission* with the GIVE, LET/ALLOW and PERMIT verbs of Modern Irish (Nolan 2013ab, 2012ab, 2011, 2010) and associated issues with syntactic construction patterns, argument realisation, event chaining (Talmy 2000) and complex multi-verb clauses. We explore these concepts as part of the dimensions of causation and causative purpose, concentrating on indirect causation, and consider other related factors such as control over causer/causee in causal event chaining in complex clauses.

Talmy's notion of force dynamics characterises causation in terms of the impact of force dynamics over causal event chains. As such, the event-frame evokes a set of conceptual elements and inter-relationships, which are felt to be central and to constitute a coherent unit with multiple event chains represented in syntax over different temporal extents, some close together and some not so close in time. The causal chain is initiated by the agent's volitional act and this, with the goal, respectively marks the beginning and end of the causal chain event-frame.

Across these event frames, capturing a causing outer event and a caused inner event, the syntactic pattern of realisation of shared arguments over these multi-verb, multi-event constructions with various argument structures is best characterised within the nexus juncture relation of co-subordination. In such complex constructions the multiple events are embedded and arguments shared across verbs in nexus juncture relations (Nolan 2012, Van Valin 2005) and we characterise these for Modern Irish. The paper addresses the mapping at the semanticsyntactic interface across these verbs in the constructions (Nolan, Rawoens and Diedrichsen 'to appear 2014', Nolan and Diedrichsen 2013, Nolan and Perrián 2014) in which they occur and the account is presented within a functional characterisation.

Causative *lassen* constructions in German: Syntax, argument structure, meaning variants and the impact of cultural knowledge in disambiguation

Elke Diedrichsen, Computational Functional Linguistics Research Group
Institute of Technology Blanchardstown, Dublin IRELAND

In many languages, causativity is a grammatical category in that there is a morphological marker or a lexical expression for a situation in which an agent brings about a state of affairs or makes somebody else do something. In German, both variants of formal marking for causativity are found only very rarely. The main expression of causativity in German is the *lassen* construction. This is a periphrastic construction with the auxiliary *lassen* ('let'), whose semantics can vary within a spectrum of meanings involving causativity, but also including permission and non-intervention.

According to Dixon (2000:30), "a causative construction involves the specification of an additional argument, a causer, onto a basic clause." The causer can be someone or something, like an event or state, "that initiates or controls the activity". A causer will therefore always appear as the syntactic relation A (actor). If a causer is added to a transitive sentence (usually expressing that X makes Y do Z), a second agentive argument is introduced to the sentence structure. In German, the causee in a causative transitive sentence is realized as an O argument with accusative case; this results in a sentence with two accusative arguments.

(1) *lassen causative with a permissive meaning*

Er lässt die Kinder den Film sehen.
3MsgNOM let.PRES3sg DEFplACC child.pl DEFMsgACC film.sg watch.INF
He lets the children watch the movie.

(2) *lassen causative for situations of tuition and instruction (intermediate between permission and direction)*

Er lässt die Kinder das Gedicht lernen.
3MsgNOM let.PRES3sg DEFplACC child.pl DEFMsgACC poem.sg learn.INF
He lets the children learn the poem = He makes the children learn the poem

(3) *lassen causative with a directive meaning*

Er lässt den Gärtner den Rasen mähen.
3MsgNOM let.PRES3sg DEFMsgACC gardener DEFMsgACC lawn.sg mow.INF
He lets the gardener mow the lawn.

Most *lassen* constructions allow more than one reading. The differences in interpretation concern the degree of control or causative force exhibited by the A argument of *lassen* (henceforth: the causer). The relationship between the causer and the event expressed may vary between the extremes of non-intervention, on the one hand, and direct and immediate causation, on the other hand. There are intermediate degrees of causativity which involve permission, for example, and direction in situations of tuition and instruction, where the causee rather than the causer has the resulting benefit. The degrees of causativity can be represented in a scale of binary features of causativity, which will be discussed in the paper.

How is the construction disambiguated in language use? In many cases, the context makes clear which degree of causativity is attributed to the causer. Also, it will be argued that many of the causative *lassen* constructions can be disambiguated on the basis of cultural knowledge (Diedrichsen 2013a, to appear), for example knowledge about the distribution of roles between people in terms of authority and responsibility. This cultural knowledge results from experience with typical situations of the kind expressed in the respective sentence.

The syntactic and semantic features of the *lassen* construction will be represented in a Constructional Schema (Diedrichsen 2013b, Van Valin 2005), which will make reference to contextual and cultural influences that help disambiguate the construction in language use.

The integration of transitive predicates into intransitive constructions: the case of the intransitive motion constructon.

Aneider Iza Erviti
Universidad de La Rioja

Goldberg (1995) postulates the existence of the intransitive motion configuration but she does not analyze it in detail. In her analysis, this configuration, which accounts for intransitive sentences expressing motion, is a subpart of the caused-motion construction, and its most distinguishing characteristic is that the external cause of the movement is not present in its formal expression. Other constructions such as the caused-motion construction have already been analyzed in detail by construction grammarians (Goldberg 1995; Goldberg and Jackendoff 2004; Ruiz de Mendoza and Mairal 2007, 2008a). However, the number of studies that deal with the intransitive-motion construction is much smaller, and if considered in relation to the integration of transitive verbs into this intransitive construction, previous publications are scarce. The present study attempts to show that the complexity of the intransitive-motion construction has been largely underestimated, and that the cognitive operations underlying this type of form-meaning pairing should be further explored under the light of a broader corpus of data. In order to achieve this aim, this presentation addresses the lexical-constructional integration of contact-by-impact verbs into the intransitive motion construction (e.g. *The fabric cracks and bangs in the wind; Jinny heard a car come bumping up*).

There have been several classifications of these verbs: Levin, in her work *English Verb Classes and Alternations* (Levin 1993: 148-156), identifies four sub-classes of verbs under the heading “Verbs of Contact by impact”, namely *Hit* verbs, *Swat* verbs, *Spank* verbs and *Non-Agentive* verbs of contact-by-impact, a total of 60. By contrast, Faber and Mairal’s (1999) lexematic-oriented classification labels these groups of verbs as “Verbs of Contact”, and instead of classifying them into four categories, they list them according to lexically-realized contextual parameters (e.g. type of blow, instrument used, entity affected, and the reason for the blow).

In my presentation I will propose a new taxonomy that overcomes the weaknesses of previous classifications of the verbs under study and I will discuss the constructional use of these verbs on the basis of real language data from three different corpora: the *British National Corpus* (BNC), the *Contemporary Corpus of American English* (COCA) and the *WebCorps*. This study also depends mainly on the *Collins Cobuild Dictionary* (1987) for the definition and delimitation of the senses of the predicates included in the analysis. This dictionary reveals the most salient semantic and pragmatic properties of each of the verbs, which are necessary for an account of the factors that license the integration of a particular verb into a given construction. Finally, I will explain the compatibility of such verbs with the intransitive-motion construction by using the explanatory tools provided by the *Lexical Constructional Model* (Ruiz de Mendoza and Mairal, 2008a, 2008b, 2011), especially the set of *internal* and *external constraints* involved in lexical-constructional integration.

The main conclusions that derive from this study are: (1) most contact-by-impact verbs are compatible with the intransitive-motion construction and that this integration is possible by applying different high-level metaphors and metonymies such as A COMMUNICATIVE ACTION IS AN EFFECTUAL ACTION (*We must hit back with the truth*), CHANGES OF STATE ARE CHANGES OF LOCATION or CONTROLLER FOR CONTROLLED (*Bob...kicked forward into an accelerating gallop*); (2), all contact-by-impact verbs that are related to sound (*batter, bump, crack, etc.*) are licensed into the intransitive-motion construction through the application of the high-level metonymy RESULT FOR ACTION; the verb merges the action performed by the subject with the sound produced by that action, and the resulting sound is taken as the action performed by the subject (*Her head went down sharply, cracking against the top of the wall*).

Towards the meaning and realization of Māori neuter verbs

Aoife Finn

Trinity College Dublin

Traditional Māori grammars recognise up to five verbal classes according to Harlow (2007: 104). These include the neuter verb class and the transitive verb class. The Māori subject is unmarked and follows the verb. As expected, Māori transitive verbs have two arguments, see example (1) from Bauer (1997: 40). In a transitive verb construction, the actor is realized as the subject, or the privileged syntactic argument in Role and Reference Grammar, see Van Valin (2005). In contrast neuter verbs have only one obligatory argument, see (2) from Harlow (2001: 31). This neuter verb subject, the one obligatory argument, is always an undergoer. If desired an actor may be expressed by a preposition *i* as in example (3), (ibid). The exact status of the actor in a neuter verb construction is unclear. While neuter verbs have only one obligatory syntactic argument, they differ from the intransitive class and the so-called passive voice. Furthermore, neuter verbs behave differently with respect to a number of constructions including nominalizations and the so-called passive voice.

(1) E kai ana ngā tamariki i ngā āporo TNS eat TNS DET children ACC DET.PL apple The children are eating the apples

(2) I pakaru te wini TNS break DET window “The window was broken”

(3) I pakaru te wini i a Tamahae TNS break DET window AGT ART Tamahae “The window was broken by Tamahae”

(4) Ka hopu ngā whānua o Ngāi Tahu i te tītū mō te hoko TNS catch Ngāi Tahu families ACC DET bird for DET sale “Muttonbirds were harvested and sold by Ngāi Tahu families”

(5) Kua mau i a Tū tētahi manu TNS catch AGT ART Tū DET bird “Tū has caught a bird”

It is often the case that a state-of-affairs may be described using either a neuter verb or a transitive verb.

This is the case in examples (4) and (5). *Te Ara* provides example (4) which includes the transitive verb *hopu* and it is glossed as ‘catch’. Example (5), taken from Bauer (1993: 413), includes the neuter verb *mau* and it is also translated as ‘catch’. It seems that neuter verbs and transitive verbs have similar, if the same, meanings. Yet the syntactic realization of the neuter verbs and transitive verbs is quite different.

Māori exhibits both ergative-like and accusative-like constructions making correct categorization difficult. Given their undergoer subject, neuter verbs are often considered as evidence of ergativity in Māori. Biggs (1969: 56) regarded them as stative verbs. Therefore, careful consideration of the neuter verbal class is an essential step in the correct categorization of Māori. Given their different syntactic realizations, this paper preliminarily considers if the meaning of transitive and neuter verbs are in fact the same. The tests for determining predicate classes tests in Role and Reference Grammar, see Van Valin (2005: 31-50), provide a means to assess the logical structure of neuter verbs. Unaccusativity, as discussed by Levin (1995), is also briefly discussed concerning neuter verbs.

A systematic approach to labelling external constraining factors in the lexical constructional model

M^a Sandra Peña Cervel
University of La Rioja
sandra.pena@unirioja.es

This proposal provides a new way of labelling the high-level metaphors and metonymies which act as external constraining factors in the lexical-constructional process of fusion within the framework of the Lexical Constructional Model, a theoretical paradigm devised by Ruiz de Mendoza and Mairal (Ruiz de Mendoza 2013; Ruiz de Mendoza and Mairal 2008, 2011). These authors and their followers are neither systematic nor consistent in their labelling system of external constraining factors. In the pioneering works by Ruiz de Mendoza and Mairal (2008, 2011), these scholars state that the notions of ‘effector’ (the argument which does something) and ‘effecting’ (which means ‘acting’) are inspired in terminology used in connection with functional approaches that contemplate semantic functions or roles as related to Aktionsart characterizations. Finally, Ruiz de Mendoza and Mairal (2008) coin the term ‘effectee’ (the argument which represents the affected entity). However, in subsequent works by Ruiz de Mendoza, Mairal and their collaborators they do not devote any section to clarifying the details on the system of labelling they will use in their analyses of the cognitive mechanisms which license lexical-constructional subsumption or fusion. In their works, we find a clear functional orientation by resorting to labels derived from the Aktionsarts distinguished by Dik (1997) and Halliday (1994) (for example, in the high-level metaphor AN ACTIVITY IS AN EFFECTUAL ACTION or in the high-level metonymy INSTRUMENT FOR ACTION). On other occasions, as is the case with GETTING RID OF A PROPERTY IS GETTING RID OF A MOVING OBJECT the labels are made up *ad hoc*, serving a particular purpose or need. We take sides with this idea since new metaphors and metonymies have to be discovered in order to be able to analyze the intricacies of language but they should conform to a functional, cognitive or any other orientation in order to provide a systematic list of high-level metaphors and metonymies which belong to an open set which is ready to accommodate new items as research develops through time. Finally, especially if we pay attention to Baicchi’s analysis of the caused-motion construction in terms of its underlying high-level metaphors, we realize that she also uses labels which correspond to Role and Reference Grammar (accomplishments for instance) and to Halliday’s Functional Grammar (for instance mental acts are what he calls mental processes).

These works devised within the Lexical Constructional Model have gone a long way towards elucidating many motivational devices of sets of verbs which participate in certain constructions. However, we would like to offer a more systematic way of providing labels for the high-level metaphors and metonymies which constrain the process of fusion of lexical items into constructions. By way of illustration, this proposal will focus on the high-level metaphors and metonymies which allow the compatibility of different sets of predicates with the fake reflexive construction. With this aim in mind, we will take as a basis Halliday and Matthiessen’s (2004) semantic classification of process types into material, behavioural, mental, verbal, relational, and existential processes. The main reason for this is that it provides many essential distinctions which help us to offer fine-grained analyses of the topic under study. We will hardly have the necessity to resort to our imagination in order to create new labels which can differ from author to author (Peña 2014). Thus subjectivity will be avoided to a great extent.

Formation of a substantive core of the polysemous verb of relations

Svetlana Kiseleva

Nelly Trofimova

This investigation is implemented within the cognitive approach and is dedicated to the formation of a substantive core of the polysemantic verb, particularly of the verb of relations (for example, the verbs with the meaning of “part and whole) in the Modern English. The cognitive approach is suspected image mainstay as well as third meaning mainstay; it means the meaning of the speech act, which is realized by the speaker in the proper communicative situation or inside of the fitting context on the basis of the nominative and primitive terms, included in the words. The substantive core of the polysemantic word is appealed to integrate all its lexical-semantic variants (LSV) and not to allow it falling into homonyms.

The substantive of the meaning of the word has the following features: narrowing of the semantic elements till min necessary, stable components, where the meaning should be max recognizable; the opinion subjectivity is admissible from the point of view of the cognitive linguistics.

Thus, it is possible to define the main cognitive mechanisms, which are the base of verb meaning «compose», and to define an invariant as the substantive core, connecting LSV of this verb. The purpose is either to prove the functioning of the lexeme of the verb representative, expressing the relations between the part and the whole on the level of the language system and updating of transferred meanings at the speech level.

In the capacity of LSV of the researched verb, motivated by the nominative-primitive meaning, it is reasonable to make an example of any secondary meaning of this verb. Its analysis is realized on the basis of comparison as the traditional way of interpretation. It is believed an actual thing to find out the items are the basis while the formation and decoding of the metaphorical statements. The analysis is based on the principles of the cognitive approach, supported of cognition and nomination for the appropriate images of perception. It is in prospect to prove whether the cognitive images are kept while understanding of the metaphorical statement. The cognitive approach supposes the image support as well as the third meaning.

The lexeme ‘compose’ keeps the meaning of establishment of the partial relations between ‘the whole’ and its ‘parts’: ‘the whole’ and ‘parts’ are represented explicitly. So, when the verb ‘compose’ has secondary meanings, the nominative-primitive meaning is seen through the substantive core of this verb. In these metaphors there is an adaptation of the verb meaning ‘to compose’, implied of the direct meaning, but kept the idea to ‘combine the parts into the whole’. At such abstract level the conceptual integration happens.

The analysis results allow us to form the following core of the lexeme ‘compose’ as ‘to acquire integrity’. This invariant as the substantive core is an item of the lexical language system and all meanings are updated against the communicative set of participants of communication.

In conclusion, it should be noticed that the abstract meaning ‘to acquire integrity’ is a strong reason of real existing of the substantive core connecting all verbs of one class but not showing the different features of separately taken verb.

Can Conceptual Integration Optimise the Activation of Embedded Scripts?

Alexey Yavetskiy

Implementing any natural language processing theory in practice means, for instance, developing efficient procedural scenarios which would ensure a quick and efficient meaning analysis/synthesis. Such procedural scenarios are responsible for activating embedded scripts (such as WORKING WITH CASH which is embedded in EATING OUT, see Periñán-Pascual 2012) and, more generally, they provide the speaker with rules for obtaining inferential knowledge.

To provide a full-fledged account of complex scripts, it is necessary to model the cognitive processes which underlie the activation of embedded scripts. To this end, the first part of the paper outlines the approach to embedded scripts as adopted in FunGramKB, which combines the principles of Role and Reference Grammar (Van Valin 2005) and those of situated conceptualisation (Barsalou 2003, 2009). According to this approach, embedded scripts are activated because the semantic roles in the host script are instantiated by the same participant (i.e., are filled by the same lexical item, CASH) as in the guest script (Periñán-Pascual 2012).

Thus, the concept of embedded scripts is based on the idea that a certain propositional element of a host script is mapped onto another script which may be “unfolded” if so required by the language analyser. However, in order to understand how the human mind is able to process language quickly and efficiently, we should examine the details of such mappings. In the second part of the paper, mappings of embedded scripts are approached from the viewpoint of Conceptual Integration Theory (CIT) (Fauconnier 1997, Fauconnier & Turner 1998). It is demonstrated that mappings of embedded scripts match CIT’s methodological concept of cross-space mappings. Cross-space mappings, for their part, occur owing to (a) the semantic participants of the analysed lexical item and (b) the semantic class to which this item belongs. Thus, EATING OUT implies an agent, some location to which the agent goes, and the catering service that the agent receives at that location. This allows us to categorise the script EATING OUT as a member of the class RECEIVING SERVICES (a type of ACTIVITIES). However, a more fine-grained analysis shows that some of the components of EATING OUT (and any script of that kind) represent complex conceptual structures themselves, or, in terms of CIT, conceptual blends. Contrary to embedded scripts, such blends (which we call “embedded mental spaces”) may account for the conceptual complexity (a) of the settings and (b) the participants of the scripts. We argue that such blends, rather than being pre-assigned to scripts, are (re)constructed based on the following procedural scenario: if there are similar thematic roles within different components of a script, then the lexical item instantiating a thematic role in the host script is most likely to trigger cross-space mappings of the similar thematic roles in the embedded scripts (e.g., the Recipient of catering services in the host script EATING OUT maps onto the Recipient of food, Recipient of change, etc.).

Interpretation of co-referential chains in Czech

Alena Poncarová

My research considered the methods of constructing the meaning of the co-referential chains in Czech and factors influencing it.

I tested contexts of two sentences under the several conditions: The first sentence contained two noun phrases. The second one contained a pronoun (possessive or personal). There were no preferences for neither of the two noun phrases to be pronoun's antecedent. An example of the test context:

(1) *Tomáš pozval na zápas Michala. Je velký fanoušek fotbalu.*

Thomas_T,Sub invited to match Michael_F,Obj. Is huge fan soccer.

Thomas invited Michael to the match. He is a huge fan of soccer.

(2) *Michala pozval na zápas Tomáš. Je velký fanoušek fotbalu.*

Michael_T,Obj invited to match Thomas_F,Sub. Is huge fan soccer.

Thomas invited Michael to the match. He is a huge fan of soccer.

(3) *Tomášem byl na zápas pozván Michal. Je velký fanoušek fotbalu.*

Thomas_T,Obj was to match invited Michael_F,Sub. Is huge fan soccer.

Michael was invited to the match by Thomas. He is a huge fan of soccer.

I controled two possible factors: the information (topic-focus articulation) and the constituent structure (subject-object function). My hypotheses were:

(A) It is more often to choose the noun phrase within the topic position as an antecedent of the pronoun.

(B) It is more often to choose the noun phrase within the subject position as an antecedent of the pronoun.

Almost 300 respondents participated on the survey and have chosen the appropriate antecedent for the pronoun in the time limit of 20 seconds. The results show that the information and constituent structures were the key factors; both hypotheses were, as whole, confirmed. Following table shows the results for a few contexts, for example in the column „T“: respondent has chosen the noun phrase in the topic position as an antecedent under the topic-condition 65 times and under the focus-condition 5 times.

contexts	T	F	T-F	Sub	Obj	Sub-Obj
1	65 5	14 55	p < 0.05	62 8	20 49	p < 0.05
2	60 11	10 58	p < 0.05	56 15	45 25	p = 0.083
3	25 46	32 38	p = 0.272	32 39	61 9	p < 0.05
4	21 49	39 28	p < 0.05	25 45	58 12	p < 0.05
5	37 34	48 23	p = 0.87	37 34	52 19	p < 0.05
6	68 3	2 69	p < 0.05	64 3	10 59	p < 0.05
7	40 31	34 35	p = 0.504	46 23	44 27	p = 0.687

Of course, there are additional factors to be considered: concrete lexical cast, respondent's individual experience or the fact that the co-referential unspecificity on the language level is quite common in the communication.

Automatic coreference resolution with fungramkb

María José Ruiz Frutos M.A.

Bayreuth University

mjruiuz@gmx.de

Today, Natural Language Processing systems are facing the difficulty of appropriately interpreting *anaphors* (Halliday and Hasan, 1976; Hirst, 1981; Kuno, 1987) in texts and rightly connecting *coreferences*, i.e. two elements that refer to the same extra-linguistic referent in the real world. Different projects have examined the possibilities of including world knowledge, common sense knowledge, and/or deep semantics in order to emulate human reasoning in (a) the process of *anaphora resolution*, i.e. the ability to determine the right antecedent (cf. Mitkov, 2002); and in (b) the process of *coreference resolution*, i.e. the ability to find the right referent in the real world by resolving different denominations for the same entity (cf. Hendrickx, Hoste and Daelemans, 2007). And this is exactly why the knowledge base FunGramKB is being suggested: it will analyze how the automatic coreference resolution can benefit from the knowledge stored in FunGramKB (cf. Perrián Pascual and Arcas Túnez, 2010); and not only the cultural information stored in the Onomasticon, but also the procedural information in the Cognicon and the semantic information in the Ontology.

The frame of this paper will be specifically the coreferential elements in the shape of lexical noun phrase anaphora for both for English and Spanish. *Indirect anaphors* (i.e. anaphors that require world knowledge in order to be resolved, cf. Hendrickx, De Clercq and Hoste, 2011) will also be commented upon, since their resolution is a real challenge due to the semantic associations that must be taken into consideration.

After examining in a theoretical way which knowledge stored in FunGramKB can contribute to successfully resolving coreferences, the importance of the reasoning performed in this knowledge base will be pointed out: not only inheritance, inference (cf. Perrián Pascual and Arcas Túnez, 2007), and internal script activators (cfr. Perrián Pascual, 2012); but also MicroKnowing, MacroKnowing (cf. Perrián Pascual and Arcas Túnez, 2005) and a conceptual-knowledge spreading based on the Onomasticon (cf. Perrián Pascual and Carrión Varela, 2011).

Finally, contributions of semantic web technologies for the automatic coreference resolution will be mentioned.

The contribution of FunGramKB to study semantic memory loss

María Beatriz Pérez Cabello de Alba (UNED)

bperez-cabello@flog.uned.es

Eva Samaniego Fernández (UNED)

esamaniego@flog.uned.es

Dementia disorders present loss of cognitive functions. Peraita and Grasso (2010a,b) have conducted studies on semantic categories in healthy individuals and patients with Alzheimer to analyse the conceptual system damage in the latter. The authors propose a model based on empirical data, which concludes that a healthy subject will be able to verbalize a category with a number of characteristics based on very general relations such as evaluation, taxonomy or function.

Our proposal is to enrich the studies of semantic memory using FunGramKB ontology (Periñán y Arcas 2004, 2005, 2006, 2007). We start from the premise that the two types of long-term memory distinguished by Cognitive Psychology, namely episodic memory and semantic memory, are correlated in FunGramKB with the Onomasticon and Ontology, and the Cognicon, respectively. Thus, the concepts of the ontology in FunGramKB would account for the categories, and the predications that define each concept would enrich the characteristics of the categories substantially. Therefore, since studies on conceptual system impairment have concluded that there are semantic category-specific deficits, i.e. not all concepts are affected in the same way, through the networks established between concepts in FunGramKB we can identify recoverable or lost cognitive strips and contribute to a more detailed diagnosis of semantic memory loss.

An analysis of the prototypical senses of prepositions *in*, *on* and *at* to be implemented in an NLP system

Diana Hernández Pastor

The relational function of prepositions when combining with other grammatical categories triggers a high degree of ambiguity in their semantic description. In an attempt to explain the polysemous behaviour of prepositions, special attention has been paid to the search for a prototypical sense of these lexical items. Outstanding works by Lakoff (1987), Herskovits (1997), Evans and Green (2006) or Taylor (2009) have contributed to the theoretical orientations that try to account for how through semantic extension some relational conceptualizations can help clarify the varied senses of the same word. For this reason, and with the aim of proceeding to the construction of an NLP semantic repository that allows the predication of prepositions within the framework of FunGramKB and the Role and Reference Grammar, this paper proposes a qualitative analysis of some examples of the prototypical senses of *in*, *on* and *at* extracted from a corpus that may serve as an attempt to formalize semantic data of these prepositions in the parser Artemis.

An account of English interrogative structures within ARTEMIS

María Auxiliadora Martín Díaz, Eulalia Sosa Acevedo
y Francisco José Cortés Rodríguez
mmartind@ull.edu.es; eusosa@ull.edu.es; fcortes@ull.edu.es
Instituto Universitario de Lingüística Andrés Bello
Universidad de La Laguna

As stated in Periñán-Pascual and Arcas Túnez (2010: 2671) FunGramKB's lexical and grammatical levels are grounded in sound linguistic theories, allowing the system to capture syntactic-semantic generalizations which are able to provide both explanations and predictions of language phenomena. They further comment that the linguistic foundation of FunGramKB is inspired on Role and Reference Grammar (RRG) and the Lexical Constructional Model (LCM). The close interrelation between these grammatical theories and FungramKB is patent in the design of ARTEMIS (Automatically Representing Text Meaning via an Intelingua-Based System"), an NLP system whose objective is the simulation of natural language understanding. In doing so, ARTEMIS is a development of the syntax-to-semantics linking algorithm proposed in RRG. This involves a parsing process which starts with a natural language sentence, extracts its morphosyntactic features and provides a representation of these in terms of the so-called layered structure of the clause in RRG.

Despite its centrality, there is a need for a full development of ARTEMIS. This presentation, therefore, aims at contributing to its development by proposing the set of rules and operations necessary for the treatment of interrogative sentences within such an NLP system.

ARTEMIS: Designing the rules for *do* insertion in English

Ana Díaz Galán and María del Carmen Fumero Pérez
mfumero@ull.edu.es; adiazgal@ull.edu.es
Instituto Universitario de Lingüística Andrés Bello
Universidad de La Laguna

The Functional Grammar Knowledge Base -FunGramKB- (Periñán Pascual y Arcas Túnez 2010) has been designed to be used in different Natural Language Processing tasks, such as information retrieval and extraction, machine translation, dialogue based systems (question-answering), etc. In order to comply with any of these tasks, it is necessary to complement the knowledge base with a device that binds natural language fragments with their corresponding grammatical and semantic structures. ARTEMIS (Automatically Representing Text Meaning via an Interlingua-Based System) has been developed with this aim. This prototype consists of three components: The G(rammar)D(velopment) E(nvironment), the CLS Constructor and the COREL-Scheme Builder (Periñán-Pascual y Arcas Túnez 2014). The GDE comprises the set of grammatical rules necessary for the parsing of natural language expressions; such rules yield, as a result, a parsed tree following the format, with some modifications, of the Layered Structure of the Clause as proposed in Role and Reference Grammar.

This work offers a partial view of these rules since we will focus specifically on the rules which govern the morphosyntactic features that trigger the insertion of the auxiliary element 'do' in English structures.

The integration of the concept +CRIME_00 in FunGramKB and the conceptualization or hierarchization problems involved.

Ángela Alameda Hernández
Universidad de Granada
aalameda@ugr.es

María Enriqueta Cortés de los Ríos
Universidad de Almería
mecortes@ual.es

Ángel Felices Lago
Universidad de Granada
afelices@ugr.es

In the last few years, a few contributions have been made in an attempt to connect the Core Ontology of FunGramKB (Periñán and Arcas, 2010) to other domain-specific ontologies (Ureña Gómez-Moreno, Alameda Hernández, Felices Lago, 2011; Felices Lago and Ureña Gómez-Moreno, 2012, 2014; Faber, Mairal and Magaña, 2011; Periñán and Arcas, 2014). The main purpose of these studies is to expand the lexical and conceptual repositories integrated in FunGramKB and facilitate the application of NLP tasks to expert knowledge.

In this paper we focus our attention on the research conducted so far in the *Globalcrimeterm* Satellite Ontology (integrated in FunGramKB) and, more precisely, on the methodological problems encountered in the conceptualization of the superordinate basic concept +CRIME_00 and its configuration as an umbrella concept. To start with, +CRIME_00 is included in the subontology of Entities (Core Ontology) as a mirror concept and is itself a superordinate entity that comprises a variety of other terminal concepts whose meaning postulates have been properly defined and organized. The process has followed the COHERENT methodology as described by Periñán and Mairal (2011), comprising the phases of conceptualization, hierarchisation, remodeling and refinement. The present paper illustrates this process with a sample of +CRIME_00 subordinate concepts such as \$ASSAULT_00, \$ABDUCTION_00, \$TRAFFICKING_00, \$CYBERCRIME_00 and \$LAYERING_00, whose meaning postulates have been created using the metalanguage known as COREL (Periñán and Mairal, 2010). The selection of the conceptual units has been assisted by the compilation of a specialized corpus (semi-automatic process) and followed by a thorough lexicographical analysis of general and specialized sources (manual process). This double-check method has helped the researcher to determine whether the selected concepts belong to the general FunGramKB Core Ontology or whether they should be included in the domain-specific ontology (*Globalcrimeterm*).

The population of lexical units corresponding to *Globalcrimeterm* concepts in the FunGramKB's lexico-grammatical module

María Diana Fernández Lloret
University of Granada
Granada (Spain)
dianallore@gmail.com

María Ángeles Gómez Castejón
UNED, Madrid (Spain)
angeles.gomezcastejon@hotmail.com

There is a constant need for specialized knowledge organization and accessibility for Natural Language Processing purposes, Knowledge Bases and Artificial Intelligence systems and, among other valid alternatives, FunGramKB (a multilingual, multipurpose knowledge base rooted in a deep semantics approach) which could play a leading role in this context. In the present paper, we will focus on some aspects of the development of the lexico-grammatical module of FunGramKB as an interface between syntax and semantics and applied to a domain-specific field under construction.

In this sense, most theories on lexical knowledge representation present a collection of grammatically salient characteristics as the unique components in a lexical representation theory, on the grounds that only those aspects which are grammatically relevant serve as the input for grammatical processes. As an example of analysis of different types of verbs, Pesetsky (1995: 14) and Grimshaw (1993: 3), in their studies of speech act verbs and verbs of color respectively, stated that the parameters “loud” and “soft” or the color parameters do not have any role in grammatical processes (Mairal, 2003). The approach presented in this paper diverges from such proposals since we believe that there are certain semantic patterns which do not actively interact in the different structural realizations, but there are others which highly limit the different syntactic configurations, e.g. the manner vs. result constants (Mairal, 2003). Consequently, in our opinion, lexical representations according to the Lexical Constructional Model (LCM) should apprehend those aspects of meaning which are grammatically relevant and should also provide a comprehensive description of the full inventory of parameters involved in meaning construction, including those related to pragmatic, semantic and discourse information. Therefore, we use the term lexical template to refer to this new lexical notational device that combines the set of grammatically, semantically and pragmatically relevant features (Mairal and Ruiz de Mendoza, 2008).

The purpose of this paper is to describe the lexical templates where lexical units are stored in the FunGramKB Suite (the lexico-grammatical module) and how each lexical unit is linked to an underlying meaning representation or concept, which is, in turn, linked to other concepts by means of a subsumption relation. The lexico-grammatical module is made up of lexical templates for the lexical units corresponding to specific entities, events and qualities, and we will enumerate their main characteristics. In order to populate the lexico-grammatical module with the appropriate lexical units, we have first selected a number of relevant concepts from FunGramKB Core Ontology or the Satellite Ontology under construction. Then, we will describe how the lexico-grammatical module is linked to the Core Ontology or the “Globalcrimeterm” satellite ontology, depending on the level of specificity of the concepts. In the “population process” we will fill in the slots and describe the main pragmatic and syntactic characteristics of the relevant lexical units, including *aktionsarten* or other complex syntactic patterns. We will conclude with statistical considerations of the results obtained and possible suggestions for a better functioning of this lexical module.

Procedural processes in Criminal Law: conceptualization of expert knowledge in *FunGramKB* Ontology¹

María de Gracia Carrión Delgado,
Universidad Nacional de Educación a Distancia
mdgcarrion@madrid.uned.es

Alba Luzondo Oyón,
Universidad Nacional de Educación a Distancia
aluzondo@flog.uned.es

FunGramKB is a knowledge base particularly designed for natural language understanding systems designed to be reused in various NLP tasks (information extraction and retrieval, machine translation, dialogue-based systems, etc). Its conceptual level comprises three main modules: the Ontology, the Cognicon and the Onomasticon. The Ontology represents a hierarchical catalogue of concepts that describe semantic knowledge in three subontologies, namely #ENTITY, #EVENT and #QUALITY. In this contribution we elaborate a thorough analysis of two concepts from the procedural criminal law domain. For this purpose, we use the COncceptual REpresentation Language (*COREL*), which serves as the metalanguage of the conceptual level of the knowledge base. The semantic power of *COREL* to represent common-sense knowledge has been widely demonstrated in Periñán and Arcas 2004, Mairal and Periñán 2009 and Periñán and R. Mairal 2010, among others. Therefore, the aim of this paper is to prove that *COREL* can also be reused to convey specialized knowledge by providing two examples from both the #ENTITIES and #EVENTS subontologies.

FunGramKB Ontology distinguishes three different conceptual levels: metaconcepts, basic concepts and terminals which are organized hierarchically. The upper level is composed by 42 metaconcepts in capital letters preceded by the “#” symbol and they represent cognitive dimensions around which the conceptual units are organized. Some metaconcepts of the #ENTITY subontology are #ABSTRACT and #PHYSICAL.

In an intermediate level we find the basic concepts preceded by the “+” symbol and followed by an underscore “_” and a numerical index (for example, +PUNISHMENT_00, +VIOLENCE_00, etc). These are used in *FunGramKB* as defining units which enable the construction of meaning postulates for basic concepts and terminals, as well as taking part as selectional preferences in thematic frames.

Finally, there are the terminal concepts preceded by the “\$” symbol and also followed by the underscore “_” and a numerical index. They represent the final nodes in the conceptual hierarchy. Examples of terminal concepts are \$ADMINISTRATIVE_DETENTION or \$ARREST_WARRANT_00.

In this paper we shall concentrate on the subontologies #ENTITY and #EVENT to provide some examples of procedural processes in criminal law. In the course of conceptualization of the specialized corpus, procedural criminal law terms appeared that include cases of derivation as the following examples show. In the case of nouns derived from verbs, the index “_D” and its corresponding numerical index was added to the event in order to avoid functional problems to the reasoner running on the ontology as well as to minimize redundancy and maximize informativeness in the semantic knowledge repository of *FunGramKB* (cf. Periñán and Arcas, 2005).

To illustrate this, we present below two examples of derivation in the above mentioned domain. One concept for the #EVENT subontology (\$ARREST_00), and another one for the #ENTITY subontology (\$ARREST_D_00). We provide the thematic frame (1), meaning postulate (2) and

¹ This research is part of a project supported by the Spanish Ministry of Economy and Competitiveness, code FFI2010-15983, 2011-2013: "Elaboración de una subontología terminológica en un contexto multilingüe (español, inglés e italiano) a partir de la base de conocimiento *FunGramKB* en el ámbito de la cooperación internacional en materia penal: terrorismo y crimen organizado".

natural language equivalent (3) for the former; and the meaning postulate (4) and natural language equivalent (5) for the latter. We should remark that the thematic frame portrays the prototypical participants involved in the cognitive situation conveyed by the event.

(1) (x1: +POLICE_00)Agent (x2: +HUMAN_00)Theme (x3)Location (x4)Origin (x5: +POLICE_STATION_00)Goal

(2) +(e1: +TAKE_01 (x1)Agent (x2)Theme (x3)Location (x4)Origin (x5)Goal (f1: (e2: +THINK_00 (x1)Theme (x6: (e3: past +DO_00 (x2)Theme (x7: +CRIME_00)Referent))Referent))Reason)

(3) The legal authorities take or keep a person in custody.

(4) +(e1: +BE_00 (x1: \$ARREST_D_00)Theme (x2: +OCCURRENCE_00)Referent) +(e2: \$ARREST_00 (x3)Agent (x4)Theme (x5)Location (x6)Origin (x7)Goal (f1: x1)Scene)

(5) When the police take someone away and guard them because they may have done something illegal.

After a brief presentation of *FunGramKB* Ontology, where the basic concepts are its main axis as they are the building blocks to construct the meaning postulates, this paper discusses the reuse of *FunGramKB* notational system to represent specialized knowledge of procedural criminal law terms.

Friday, 4 July 2014

A Polysemy Account of Turkish Spatial Noun ‘üst’ in Dative Case Marker

Aysun Balkan

Bogazici University, Istanbul, Turkey

aysun.balkan@boun.edu.tr

This paper examines the construction polysemy of the Turkish spatial noun ‘üst’ suffixated with dative case marker from a cognitive semantics perspective. ‘üst’ roughly translates into English ‘on’ and ‘over’. In this corpus study, I attempt to account for the polysemy network of Turkish ‘üst + Dative’ within the Principled Polysemy model proposed by Tyler and Evans (2001a, 2003). The current analysis of ‘üst + Dat’ construction indicated it encodes both contact and noncontact situations between a Trajector (TR) and a Landmark (LM). Despite the fact that Turkish and English are typologically distinct languages and express spatial relations using very different linguistic elements, the current study showed a surprising amount of overlap with the primary and extended senses found in the polysemy networks of “üst + Dative’ and English ‘on, over’. Overall, the results shed light on the applicability of the Principled Polysemy model, which has originally been developed for English spatial prepositions to the spatial nouns of a historically unrelated language such as Turkish.

Objects as symbols and clearly distinguishable symbols naming the same objects in previously isolated human groups before the origin of syntax. A hypothesis about origin of meta-symbols and superordinate categorization

Ciro Antunes de Medeiros

Universidade Estadual de Campinas (UNICAMP); São Paulo, Brazil.

c-a-medeiros@hotmail.com

Experiments results indicate a relationship between vocabulary that contains only count nouns and the inability to recognize superordinate categories by one year old children. These results suggest that the capacity to recognize superordinate categories is correlated to the ability to produce words from classes such as verbs and prepositions, as early as in 13 month old children. However, these results are insufficient to distinguish between factors that are causes from those that are effects (WAXMAN, 1995). The mastery of the production of superordinate categories supports the ability to include numerous clearly distinguishable items in the same category (MURPHY, 2002). By thinking in a category as a symbol, the ability to recognize superordinate categories would be a process analogous to the ability of dealing with the perspectival nature of a linguistic symbol. An object can be simultaneously a flower, a rose or a gift because the perspectival nature of the human linguistic symbols multiplies indefinitely the possibilities of use of a symbol (TOMASELLO, 1999, p. 107). The ability to recognize superordinate categories would be, therefore, a necessary step to lead to the emergence of the typical human language as well the ability to recognize superordinate categories is required in the language acquisition process. The objective of the Article is to formulate a hypothesis that can induce further investigations about the origins of the ability to produce superordinate categories in an environment whose lexical stimuli was entirely, or almost entirely, composed by nouns. The research was developed from a review of the literature regarding the ability of categorization (MURPHY, 2002), the origins of human language according the principle ontogeny recapitulates phylogeny (TOMASELLO, 1999), experiments about the relation between conceptual organization and lexical acquisition (WAXMAN, 1994; 1995) and other relevant articles to the theme. The hypothesis starts from the idea the human language preceding the syntax would be like "a more primitive language", described by Wittgenstein in *Philosophical Investigations*, in which "every word has a meaning" and the ability to recognize superordinate categories is assigned to a minor importance. In this more basal language, the predominantly represented categories would be basic level ones. However, it would be from a language based on basic level categories that the ability to include clearly distinguishable items in the same category would have developed. The necessary elements would be the ability to use objects as symbols - common in children during the acquisition of language (TOMASELLO, 1999) - and the meeting of human groups that appoint tools of the same type or similar objects with clearly distinguishable names due to previous geographic isolation. The exposure to the same object that is named with different vocalizations united to the capacity to use this object as a symbol would be the trigger to create meta-symbols and to include clearly distinguishable items in a same category.

Code-mixing and Social Identity: A Visual Approach

Jayshree Chakraborty

Department of Humanities and Social Sciences IIT Kharagpur (WB) India

shree@hss.iitkgp.ernet.in

Code-mixing is a social phenomenon representing social reality on the one hand and constructing new social meaning and reality in the mind of the users on the other. Although more predominantly it occurs in the spoken language, its impact is no less when it occurs in the written text. This can be said more assertively about the television commercials which carry their impact through various modules – speech, tone, accent, attire, visual surrounding and many more. The present research focuses on the various styles that are observed in the written text of the code-mixed TV commercials broadcast through various channels in the Indian subcontinent. In the advertisements under consideration, one of the two codes that are mixed is English and the other code is any other Indian language. One of the codes is invariably English because in India the usage of English is not motivated by functional purposes alone, but a number of other significant issues related to class and identity are also attached to it. English, in India, has always occupied a very significant position, whether as a foreign language, second language or link language. With the global revolution in the fields of communication and information technology, English has acquired altogether a different status all over the world. In the Indian subcontinent, its usage and currency have taken very innovative dimensions. Consequently, different varieties of code-mixed languages have emerged. In these code-mixed varieties, one can notice mixing of English words, phrases, expressions, tones, and even script, in different degree, with those of the Indian languages. The more interesting fact to be noticed is that these different varieties of code-mixed languages are used in different contexts with different purposes by different groups of people to achieve different objectives. A close study of the various code-mixed languages can thus reveal some interesting facts related to language and identity. With this objective, the present research makes an analysis of various code-mixed language varieties where the examples display interesting ways of mixing at the levels of script as well as tone and gestures, and the results of the analysis attest the fact that these code-mixed varieties are actually associated with different group identities, basically related to younger generation. They also indicate the informal ambience of the discourse, and similarly, the mood of the discourse participants. The data for analysis have been taken from the commercial advertisements broadcast by various channels on television. Mostly the relevant data have been taken from various Hindi channels, though some advertisements from other regional channels have also been considered for supporting or refuting some arguments as the case might have been. Since the data have been taken from television discourse contexts where various modules of discourse provide significant information spontaneously, the Critical Discourse Analysis (CDA) theory has been taken as the framework. The results of the analysis are in conformity to the poststructuralists' view that 'identity is an ongoing, interactive process rather than a fixed product'.

An Iconic and a Systematic Semantic Feature of 'Irregular' Forms in English

Elena Even-Simkin

Ben-Gurion University of the Negev, Israel

elenae@post.bgu.ac.il

There is only a limited number of Noun Plurals like *foot-feet*, *goose-geese* and Past Tense verb forms such as *sing-sang*, *win-won* that are produced by Internal Vowel Alternation (IVA) and nowadays these forms are considered to be irregular in Modern English. That is, they do not follow the rules of Past Tense /+(e)d/ and Noun Plural /+(e)s/ formation.

However, historically, these IVA forms were both more numerous and even productive in Old English and those that have remained in the language retain largely the same IVA patterns in Modern English. The recent study of these IVA phonological processes in the nominal and verbal forms revealed, first, two opposed iconic and polar systems consisting of *fronting* for Plural formation for nouns versus *backing* for Past Tense formation in verbs. Second, the recent study reveals that there are underlying systematic semantic features for these IVA forms, as well.

In this paper, I shall present an iconic phono-morphological analysis of the IVA Noun Plural and Past Tense forms in English in terms of the sign-oriented theory, in general, and the Phonology as Human Behavior, in particular. The results of this study demonstrate precisely the phonotactically-phonological quality of IVA which appears to be systematic and iconic, and may serve to distinguish between nominal versus verbal vowel alternation systems, thus making them easier to be identified by the computational language programs. These new findings may have further implications, for example, in the field of Web search engines, in the interactive computing and human-interactive robotics.

Interface of language and culture: The Mandaya phenomenon

Leorisyl D. Siarot, Ph.D.
Faculty, Davao Oriental State College of Science and Technology
Philippines
leosiarot@yahoo.com

The Philippines is composed of many tribal groups which exemplify the richness of the Filipino culture. Mandaya is just one of the many tribal groups in the Philippines and is endowed with a rich culture and language. In the study of language and culture, one is confronted with the question: what happens in the interface of Mandaya culture and language? The purpose of this grounded theory study is to develop the model on the interface of culture and language of Mandaya in Davao Oriental.

Interviews were conducted among the 20 participants in Tarragona, Manay, Caraga and Cateel. Observations were also done in the four municipalities. Recollection and reflexivity were also made to enrich the data which were subjected to open, axial and selective coding. The results were also validated through a survey questionnaire.

After a careful observance of the procedures of a grounded theory, the four final themes were formulated as principles namely: *Mandaya uses language to hand down culture and heritage; uses language as a symbol for concrete and abstract realities; lexical semantics is grounded on culture; and pragmatics is grounded on culture.* And the three and two mediary themes were drawn as the sub-principles for semantics and pragmatics respectively. These are *abstract meanings from the perceived phenomena; use words to signify concrete world; use words to signify abstract realm; implicatures in their discourse; presuppositions in their discourse; politeness is indicated is indicated in the use maxims, honorific and tone of voice; and culture sets the norms in the discourse in relation to age and gender.* Existing theories have been valuable in the discussion of the generated principles. The wisdom of these theories combined with the findings of this study provided a deep reflection of the interface of Mandaya culture and language.

In my research, I have coined precognition presupposition and imperative presupposition as additional types of presupposition which I have found to be true to Mandaya. I have also coined maxim of authority-reference as additional to Leech's Principles of Politeness.

This theoretical model will be of great help for those who desire to understand better the interface of culture and language especially of Mandaya. Although I know that this is not going to be exclusive soon to the Mandaya communities but will also be applied to other communities having similar features with that of Mandaya. The teachers and students of semantics, pragmatics, sociolinguistics and applied linguistics will benefit from this output.

Approaching illocutionary constructions within a computational environment

Alba Luzondo Oyón

Universidad Nacional de Educación a Distancia

aluzondo@flog.uned.es

Rocío Jiménez Briones

Universidad Autónoma de Madrid

rocio.jimenez@uam.es

The grammatical level of FunGramKB stores constructional schemata helping Role and Reference Grammar to build the syntax-to-semantics linking algorithm (Periñán, 2013: 209). In the Grammaticon, the four descriptive layers of the Lexical Constructional Model (LCM; Ruiz de Mendoza and Mairal, 2008, 2011; Mairal and Ruiz de Mendoza, 2009; Ruiz de Mendoza, 2013) are currently being implemented computationally. Taking the Level-3 Constructicon of FunGramKB as our focus, the aim of this paper is to discuss how illocutionary constructions (e.g. *Can You Forgive Me (XPREP)?*, *I would appreciate if you XVP*, etc.), which capture conventionalized illocutionary meaning, have been translated into the metalanguage employed in FunGramKB, i.e. COREL. In the LCM, illocutionary meaning is grounded in high-level situational cognitive models, which, in line with Ruiz de Mendoza and Baicchi (2007), can be identified by means of illocutionary scenarios (cf. Panther and Thornburg, 1998), that is, generalizations over everyday situations where people offer, forgive, apologize, etc. For example, a statement such as *I could use a glass of water* may stand as a request in the context of a request scenario based on the cultural convention according to which, if people manifest that they are negatively affected by a specific situation, other people are expected to help.

Thus far, twelve speech acts (e.g. advising, apologizing, condoling, promising, thanking, etc.) have been formalized in the Level-3 Constructicon, each of which consists of several illocutionary expression grouped under different, yet related, COREL schemes. By way of illustration, the above-mentioned speech act of requesting has been codified in FunGramKB by means of the COREL scheme in (1), which could be read as follows: “there is a speaker that requests something from the hearer and what s/he requests is for the hearer to do what the speaker wants. As a result, the hearer may help the speaker or not”:

(1) +(e1: +REQUEST_01 (x1: <SPEAKER>)Theme (x2: (e2: +DO_00 (x3: <HEARER>)Theme (x4: (e3: +WANT_00 (x1)Theme (x4)Referent))Referent))Referent (x3)Goal (f1: (e4: pos +HELP_00 (x3)Theme (x1)Referent) | (e5: pos n +HELP_00 (x3)Theme (x1)Referent))Result)

Since, as any metalanguage, COREL imposes its own semantics and syntax that need to be obeyed to produce well-formed structures (Periñán and Mairal, 2010), this paper details the challenges faced in order to provide a computational account of illocutionary constructions in FunGramKB.

**Finding a way through the population in FunGramKB
with a Latent Semantic Analysis-based tool**

María Beatriz Pérez Cabello de Alba
Ismael Iván Teomiro García

Within FunGramKB (Periñán & Arcas 2004, 2005, 2006, 2007) the protocol to implement the list of words in the Lexical module has consisted in manually inserting the lexical items related to a concept in the Ontology. The suite provides a slot for each of the languages contemplated in the knowledge Base and lexical items related to a given concept have to be chosen after a long dictionary looking-up process. In this work we propose to use Gallito (Jorge-Botana, Olmos & Escudero 2011), a Latent Semantic Analysis-based tool which, by means of a query, automatically generates a list of neighbour words. We maintain that this list of related words would be of great help to semi-automate the process of implementation of the lexical modules of FunGramKB. In order to illustrate this, we give an example of a basic concept and the list of related words extracted with the aid of Gallito.

From Terminography to Ontology: the Representation of Legal Concepts in FunGramKB

Gianluca Pontrandolfo

Department of Legal, Language, Interpreting and Translation Studies (IUSLIT) – University of Trieste

gpontrandolfo@units.it

Pedro Ureña Gómez-Moreno

Facultad de Ciencias Económicas y Empresariales

Sección de Filología Inglesa

Universidad de Granada

pedrou@ugr.es

The present paper focuses on the construction of domain-specific ontologies and the representation of specialised knowledge in legal settings (cf. Asaro et al. 2003).

It presents a contrastive analysis English-Spanish-Italian of conceptual units from the domain of labour law and, more specifically, from the subarea pertaining to the offences committed in the workplace (cf. Nairns 2008; Huarte-Mendicoa & Morales García 2004; Arena & Cui 2012).

Firstly, a classification of prototypical labour offences is shown, stemming from a conceptual analysis of the offences in the three legal cultures (Great Britain, Spain and Italy). Such preliminary examination is a crucial step while dealing with legal systems, especially if they belong to different legal traditions ('common law' in the case of Great Britain vs. 'civil law' in the case of Spain/Italy).

Secondly, a proposal is presented for the construction of terminographic templates that combine, on the one hand, the sociocognitive perspective on terminology (especially Temmerman (2000: 122)'s template) and, on the other hand, the onomasiologic approach to terminography (cf., among others, Sager 1990, Cabré 1999, Magris et al. 2002).

Thirdly, an approach is taken towards the implementation of terminology and terminographical templates within FunGramKB, a knowledge base for the computational processing of language (cf. Mairal & Perriñán 2009, Perriñán & Mairal 2010).

Some case studies will be presented based on key concepts of the domain under investigation.