7th INTERNATIONAL CONFERENCE ON MEANING AND KNOWLEDGE REPRESENTATION

MKR2018 @ ITB Dublin
4, 5 and 6 July, 2018

Institute of Technology Blanchardstown
Institiúid Teicneolaíochta Baile Bhlainséir

Conference Programme v1
&
Book of Abstracts

itb
Institute of Technology
Blanchardstown
Institiúid Teicneolaíochta
Baile Bhlainséir

LINC
Learning & Innovation Centre

FunGram
KB

lexicom

GOBIERNO
DE ESPAÑA
MINISTERIO
DE ECONOMÍA
Y COMPETITIVIDAD

IBM

Table of Contents

SCHEDULE - 4TH JULY WEDNESDAY 6

SCHEDULE - 5TH JULY THURSDAY 8

SCHEDULE - 6TH JULY FRIDAY 11

ABSTRACTS – 4TH JULY, WEDNESDAY 13

Functional neurolinguistics and clinical computing 14
Ricardo Mairal-Usón 14

Automatic domain-specific learning: Towards a methodology for ontology enrichment 16
Pedro Ureña Gómez-Moreno 16
Eva M. Mestre-Mestre 16

Locating semantic memory loss 17
María Beatriz Pérez Cabello de Alba 17
Ismael Iván Teomiro García 17

Grammatical words as structural dominants in linguistic schematization of cognitive experience 19
Irina Tolmacheva 19

Challenges for knowledge representation: Emergence in linguistic expressions and Internet Memes 20
Elke Diedrichsen 20

Referent tracking in narrative in three western desert dialects 22
Conor Pyle 22

KEYNOTE TALK 24
Dr. Elizabeth Daly 24
IBM 24

The role of previous discourse in detecting public textual cyberbullying 25
Aurelia Power 25
Antony Keane 25
Brian Nolan 25
Brian O'Neill 25

Implementing natural language understanding in an intelligent conversational agent 27
Gelmis S. Bartulis 27
Irene Murtagh 27

An experimental review on methods for Word Sense Disambiguation on Natural Language Processing 28
Fredy Núñez Torres 28

Discovering hazards via twitter for emergency management: a knowledge-based approach 30
Carlos Periñán-Pascual 30

ABSTRACTS – 5TH JULY, THURSDAY 31

Entrenchment of triconstituent English noun compounds 32
Elisabeth Huber 32

Parallels and contrasts between the approved adjectives in the ASD-STE dictionary and the adjectival concepts in FunGramKB Core Ontology 33
Ángela Alameda Hernández 33
Ángel Felices Lago 33

OVER in radiotelephony communications 35
Maria del Mar Robisco Martin 35

Mechanisms of metaphtonymy formation (based on English verbs with semantics “to separate”) 37
Svetlana Kiseleva 37
Nella Trofimova 37
Irina Rubert 37

Sampling techniques to overcome class imbalance in a cyber bullying context 38
David Colton 38

Motivating the computational phonological parameters of an Irish Sign Language avatar 39
Irene Murtagh 39

Motivating a linguistically orientated model for a conversational software agent 40
Dr Kulvinder Panesar 40

Speaker’s focus of interest as a basis of a text semantic model 41
Irina Ivanova-Mitsevich 41

From walled off Europe to walled in identity 42
Natalia Iuzefovich 42

On dominating principle of knowledge representation and meaning construction in discourse 43
Nikolay Boldyrev 43

Parsing complex sentences in ASD-STE100 within ARTEMIS 44
Marta González Orta 44
María Auxiliadora Martín Díaz 44

The syntactic parsing of ASD-STE100 adverbials in ARTEMIS 45
Francisco José Cortés Rodríguez 45
Carolina Rodríguez Juárez 45

A Sociolinguistic Corpus Based Investigation of Irish Sign Language Grammatical Classes 46
Robert Smith 46

ABSTRACTS – 6TH JULY, FRIDAY 47

How can one evaluate a conversational software agent framework? 48
Dr Kulvinder Panesar 48

Detection of cyber bullying using text mining 49
David Colton 49

A qualitative analysis of the wikipedia n-substate algorithm's enhancement terms 50
Kyle Goslin 50
Markus Hofmann 50

Feeding the Lexical Rules in ARTEMIS for the parsing of ASD-STE100 51
María del Carmen Fumero Pérez 51
Ana Díaz Galán 51

Functional-semantic status of lexical-grammar parenthesis-modal discourse-text «transitions» in modern English and French languages 52
Sabina Nedbailik 52

The role of persuasion processes in shaping various methods of message framing 53
Anna Kuzio 53

Metaphor-facilitated co-creation strategy in election campaigns 54
Inna Skrynnikova 54

The forms, functions and pragmatics of Irish polar question–answer interactions 55
Brian Nolan 55

A group theory for conceptual meanings (Digital Linguistics) 56
Kumon Tokumaru 56

SCHEDULE - 4TH JULY WEDNESDAY

Location: Lecture theatre @ ITB Linc

			Time		Description
			9:00-9:45		**REGISTRATION CHECK-IN & COFFEE**
			9.45-9:55	**15 min**	**CONFERENCE WELCOME**
					CONFERENCE STARTS

Session 1

1	1	**Chair: Francisco José Cortés Rodríguez**	10:00-10:30	30 min	Automatic domain-specific learning: towards a methodology for ontology enrichment **Pedro Ureña Gómez-Moreno** **Eva M. Mestre-Mestre**
	2		10:30-11:00	30 min	Locating semantic memory loss **María Beatriz Pérez Cabello de Alba** **Ismael Iván Teomiro García**
	3		**11:00-11:30**	30 min	Grammatical words as structural dominants in linguistic schematization of cognitive experience **Irina Tolmacheva**

			11:30-12:00	30 min	**COFFEE**

Session 2

2	**4**	**Chair: Fredy Núñez Torres**	12:00-12:30	30 min	Challenges for knowledge representation: Emergence in linguistic expressions and Internet Memes **Elke Diedrichsen**
	5		12:30-1:00	30 min	Referent tracking in narrative in three western desert dialects **Conor Pyle**
			1:00-2:00	1 hour	**LUNCH**
			2:00-3:00	1 Hr	**Dr. Elizabeth Daly** **IBM** **KEYNOTE TALK**
			3:00-3:30	30 min	**COFFEE**

Session 3

3	6	**Chair: Kulvinder Panesar**	3:30-4:00	30 min	The role of previous discourse in detecting public textual cyberbullying **Aurelia Power**
	7		4:00-4:30	30 min	Implementing natural language understanding in an intelligent conversational agent **Gelmis S. Bartulis** **Irene Murtagh**
	8		4:30-5:00	30 min	An experimental review on methods for word sense disambiguation on natural language processing **Fredy Núñez Torres**
	9		5:00-5:30	30 min	Discovering hazards via twitter for emergency management: a knowledge-based approach **Carlos Periñán-Pascual**

SCHEDULE - 5TH JULY THURSDAY

Location: Lecture theatre @ ITB Linc

			8:30-9:00		**COFFEE**

Session 4

4	1	**Chair: Irene Murtagh**	9:00-9:30	30 min	Entrenchment of triconstituent English noun compounds **Elisabeth Huber**
	2		9:30-10:00	30 min	Parallels and contrasts between the approved adjectives in the ASD-STE dictionary and the adjectival concepts in FunGramKB core ontology **Ángela Alameda Hernández** **Ángel Felices Lago**
	3		10:00-10:30	30 min	OVER in radiotelephony communications **Maria del Mar Robisco Martin**
	4		10:30-11:00	30 min	Mechanisms of metaphtonymy formation (based on English verbs with semantics “to separate”) **Svetlana Kiseleva,** **Nella Trofimova,** **Irina Rubert**

			11:00-11:30	30 min	**COFFEE**

Session 5

5	5	**Chair: Aurelia Power**	11:30-12:00	30 min	Sampling techniques to overcome class imbalance in a cyber bullying context **David Colton** **Markus Hofmann**
	6		12:00-12:30	30 min	Motivating the computational phonological parameters of an Irish Sign Language avatar **Irene Murtagh**
	7		12:30-1:00	30 min	Motivating a linguistically orientated model for a conversational software agent **Kulvinder Panesar**
			1:00-2:00	**1 Hr**	**LUNCH**

Session 6

6	8	**Chair: Conor Pyle**	2:00-2:30	30 min	Speaker's focus of interest as a basis of a text semantic model **Irina Ivanova-Mitsevich**
	9		2:30-3:00	30 min	From walled off Europe to walled in identity **Natalia Iuzefovich**
	10		3:00-3:30	30 min	On dominating principle of knowledge representation and meaning construction in discourse **Nikolay Boldyrev**
			3:30-4:00		**COFFEE**

Session 7

7	11	**Chair: Ángel Felices Lago**	4:00-4:30	30 min	Parsing complex sentences in ASD-STE100 within ARTEMIS **Marta González Orta** **María Auxiliadora Martín Díaz**
	12		4:30-5:00	30 min	The syntactic parsing of ASD-STE100 adverbials in Artemis **Francisco José Cortés Rodríguez** **Carolina Rodríguez Juárez**
	13		5:00-5:30	30 min	A sociolinguistic corpus-based investigation of Irish Sign Language grammatical classes **Robert Smith**
8	**4**		5:30-6:00	30 min	Feeding the lexical rules in ARTEMIS for the parsing of ASD-STE100 **María del Carmen Fumero Pérez** **Ana Díaz Galán**

SCHEDULE - 6TH JULY FRIDAY

Location: Lecture theatre @ ITB Linc

			8:30-9:00		**COFFEE**

Session 8

8	1	**Chair: Carlos Periñán-Pascual**	9:00-9:30	30 min	How can one evaluate a conversational software agent framework? **Kulvinder Panesar**
	2		9:30-10:00	30 min	Detection of cyber bullying using text mining **David Colton** **Markus Hoffmann**
	3		10:00 - 10:30	30 min	A qualitative analysis of the Wikipedia n-substate algorithm's enhancement terms **Kyle Goslin** **Markus Hofmann**
	4		10:30 - 11:00	30 min	A group theory for conceptual meanings (Digital Linguistics) **Kumon Tokumaru**

			11:00 - 11:30	**30 min**	**COFFEE**

Session 9

9	5	**Chair: Kyle Goslin**	11:30 - 12:00	30 min	Functional-semantic status of lexical-grammar parenthesis-modal discourse-text «transitions» in modern English and French languages **Sabina Nedbailik**
	6		12:00 - 12:30	30 min	The forms, functions and pragmatics of Irish polar question–answer interactions **Brian Nolan**
	7		12:30 -1:00	30 min	Metaphor-facilitated co-creation strategy in election campaigns **Inna Skrynnikova**

			1:00 - 3:00		**LUNCH** **COFFEE & CONVERSATION** **CONFERENCE ENDS**

ABSTRACTS – 4TH JULY, WEDNESDAY

Functional neurolinguistics and clinical computing

Ricardo Mairal-Usón
Universidad Nacional de Educación a Distancia, Madrid, Spain

Drawing on the initial work of Mairal (2017), the aim of this talk is to present an updated picture of the research agenda of what has been termed the Functional Neurolinguistics program which is being developed by Seconds and FunGramKB groups. The focus of this program is on the human brain and more specifically the linguistic deficits associated to the presence of either neurodegenerative diseases or brain tumors located in the linguistic eloquent area at the left hemisphere of the brain.

Firstly, in the case of neurodegenerative diseases, namely Alzheimer and Parkinson, mild cognitive impairment together with a gradual semantic memory loss is one of the most notorious manifestations (Boschy et al., 2017; Mairal and Pérez, 2017; Mortamaisa et al., 2017; Mueller et. al., 2016; Pérez Cabello de Alba, 2017). In connection with this and based on our previous research in the area of theoretical linguistics (e.g. the work on the Lexical Constructional Model, Ruiz de Mendoza and Mairal, 2008; Ruiz de Mendoza and Galera, 2014, Ruiz de Mendoza, 2017 etc.) and in the framework of computational linguistics and text mining (see DAMIEN in the FUNK Lab project at www.fungramkb.com), I would like to raise the following issues:

a) Can semantic memory loss be quantified? Can we provide a fine grained analysis of this gradual loss?
b) Moreover, can this gradual decline be mapped and correlated with *fMRI* (*functional magnetic resonance imaging*)?
c) Can we prevent and early diagnose semantic memory deficits?

As a first experiment, with the aim of providing a qualitative approach as to the process of semantic memory loss, the responses of Parkinson patients to the Hayling test have been analyzed using text mining techniques (DAMIEN). Interestingly enough, there seem to be a number of regular patterns that illustrate a more fine-grained picture than the all-or-nothing evidence obtained in a quantitative approach.

The second area of research, that related to oncological brain tumor resection in eloquent areas of the brain (Barcia et. al., 2012; Rivero-Rivero et. al. 2016; Duffau, 2017), opens a very stimulating research horizon for linguists since this involves addressing the fascinating topic of the plastic nature of the brain. In this regard, the following questions constitute our focus:

d) In the area of brain tumor resection, if it is possible to replicate the linguistic capacities in the right hemisphere of the brain so that patients after surgery can speak, what is the format of this new reinvented or rediscovered linguistic module?
e) What type of compensatory mechanisms can we provide so that the patient does not lose his linguistic capacity after a tumor has been resected?
f) In our endeavor to replicate what is potentially lost in one of the brain hemispheres, can we ascertain any differences between a monolingual and a bilingual brain or else between a child and an adult, both with a brain tumor?

The answers to these research questions are the core of and constitute the first step towards a more ambitious research program in Neuroscience using the linguistic and computational tools developed in previous projects, namely Lexicom and FunGramKB, together with neuroimaging studies (fMRI: *functional magnetic resonance images*).

References

Barcia, J.A. et. al (2012) "High-frequency cortical subdural stimulation enhanced plasticity in surgery of a tumor in Brocas area". Neuroreport: Volume 23 - Issue 5 – pp. 304–309.
Boschi, V., E. Catricalà, M. Consonni, C. Chesi, A. Moro and S. F. Cappa (2017). "Connected Speech in Neurodegenerative Language Disorders: A Review" Frontiers in Psychology, 8, Article 269.
Duffau, H. (2017) "The error of Broca: From the traditional localizationist concept to a connectomal anatomy of human brain." Journal of Chemical Neuroanatomy http://dx.doi.org/10.1016/j.jchemneu.2017.04.003
Mairal-Usón, R. (2017) "Constructing meaning for Clinical and Computational Purposes" 6th International Conference on Meaning and Knowledge Representation. St. Petersburg.

Mairal-Usón, R. and B. Pérez Cabello de Alba (2017) "Quantifying semantic memory loss: the search for linguistic compensatory mechanisms." Unpublished Technical Report.

Mortamais, M., J. A. Ash, J. Harrison, J. Kaye, J. Kramer, Ch. Randolph, C. Pose, B. Albala, M. Ropacki, C. W. Ritchie and K. Ritchie. (2017) "Detecting cognitive changes in preclinical Alzheimer's disease: A review of its feasibility". Alzheimer's & Dementia 13 (2017) 468-492.

Mueller, K. D. R., L. Koscik, L. S. Turkstra, S. K. Riedeman, A. LaRue, L. R. Clark, B. Hermann, M. A. Sager, and S. C. Johnson. (2016) "Connected Language in Late Middle-Aged Adults at Risk for Alzheimer's Disease". J Alzheimers Dis. 2016 October 18; 54(4): 1539–1550.

Pérez-Cabello de Alba, M. B. (2017) "A contribution of Natural Language Processing to the study of semantic memory loss in patients with Alzheimer´s disease". Revista de lenguas para fines específicos. Special Issue: New Insights into Meaning Construction and Knowledge Representation, 2017, Vol. 23, Núm 2, pp. 133-156.

Rivera-Rivera, P., M. Rios-Lago, S. Sánchez-Casarrubios, O. Salazar, M. Yus, M. González-Hidalgo, A. Sanz, J. Avecillas-Chasin, J. Alvarez-Linera, A. Pascual-Leone, A. Oliviero, and J. A. Barcia (2017). "Cortical plasticity catalyzed by prehabilitation enables extensive resection of brain tumors in eloquent áreas" Journal of Neurosurgery, Apr 2017 / Vol. 126 / No. 4: pp. 1323-1333.

Ruiz de Mendoza, F. (2017) "Metaphor and other cognitive operations in interaction: from basicity to complexity". In Beate Hampe (ed.) 2017. Metaphor: Embodied Cognition, and Discourse. Cambridge: Cambridge University Press; pp. 138-159.

Ruiz de Mendoza, F. and A. Galera (2014). Cognitive modeling. A linguistic perspective. Amsterdam: John Benjamins.

Ruiz de Mendoza, F. and R. Mairal-Usón (2008) 'Levels of description and constraining factors in meaning construction: an introduction to the Lexical Constructional Model'. *Folia Linguistica* 42/2 (2008), pp. 355–400.

Departamento de Filologías Extranjeras y sus Lingüísticas. UNED, 28040 Madrid, Spain.
rmairal@flog.uned.es

Automatic domain-specific learning: Towards a methodology for ontology enrichment

Pedro Ureña Gómez-Moreno
University of Granada, Granada, Spain, and
Eva M. Mestre-Mestre
Universitat Politècnica de València, València, Spain

In a world where enormous amount of data are constantly created and in which the Internet is used as the primary means for information exchange, there exists a need for tools that help processing, analyzing and using that information so that it can be properly handled. However, while the growth of information records pose many opportunities for social and scientific advance, this has also highlighted existing difficulties to process it, by extracting meaningful patterns from massive data. In this regard, ontologies have been claimed to play a major role in the processing of large-scale data, as they serve as universal models of knowledge representation, and are being proposed as possible solutions to these types of issues.

This paper presents a proof-of-concept process for ontology expansion in knowledge domains based on the exploitation of corpus and terminological data. The "ontology enrichment method" (OEM) proposed here, consists of a sequence of tasks aimed at classifying an input keyword automatically under its corresponding node within a target ontology, by following three steps: ontology identification, corpus compilation and automatic data classification.

Based on it, this paper reports on the results of a small-scale experiment carried out in the field of virology to corroborate whether the OEM was able to classify an input keyword under its corresponding superordinate in a hierarchy of viruses. Results prove that the method can be successfully applied for the automatic classification of specialized units into a reference ontology. Its main advantage is that it draws on corpus and terminological data available from academic and encyclopedic sources.

Keywords: Ontology learning, FunGramKB, Corpus, Terminology, Biology

Departament de Lingüística Aplicada. Universitat Politècnica de València. Campus de Gandia. C/ Paranimf, 1 – 46730. evamestre@upvnet.upv.es

Locating semantic memory loss

María Beatriz Pérez Cabello de Alba
Universidad Nacional de Educación a Distancia, Madrid, Spain
Ismael Iván Teomiro García
Universidad Nacional de Educación a Distancia, Madrid, Spain

Patients suffering from various types of dementia (e.g. Alzheimer's disease, mild cognitive impairment, and semantic dementia) usually present an impaired performance in several kinds of tasks concerning specific categories of objects such as animals, furniture, vegetables, etc. This ability can be selectively impaired: for example, an individual may have a compromised performance in certain tasks when living things are involved, while her performance remains relatively intact when other non-living thing categories are concerned. Hence, this impairment is known as category-specific semantic deficit, which can provide us with vital information as to how the conceptual-semantic knowledge is stored and organized in the brain (Warrington & Shallice, 1983; Basso, Capitani & Laiacona, 1988; Capitani et al., 2003, Laws et al., 2007; among many others).

Ever since Warrington & Shallice (1983), many studies have corroborated two different patterns of semantic memory loss: the dichotomy between impaired performance with living things and normal performance with non-living things, and the opposite pattern (Hillis & Camarazza, 1995). Furthermore, there are also cases that do not fit into either pattern: for example, impaired performance with living things but normal performance with body parts (Warrington & Shallice, 1984; Silveri & Gainotti, 1988), hampered performance with musical instruments (Siri et al. 2003) but not with other non-living things, and completely different patterns like the patient studied by Siri et al. (2003).

Till date, no theoretical model has been able to appropriately account for the unsystematic and varied patterns of category-specific semantic deficits found so far. Moreover, there seems to be no correlation between the type of brain damage and a pattern of memory loss, nor can the latter be accounted for by any of the so far proposed model of conceptual knowledge representation and storage in the brain (Capitani et al. 2003).

Our purpose is, thus, to adapt and provide a theoretical model that helps understand and properly interpret the available empirical data. This model is partially based on the Functional Lexematic Model (Martín Mingorance, 1998) and Peraita et al.'s (2008) model of conceptual features, and it builds on FunGramKB's ontology (Periñán Pascual & Arcas, 2007). The advantage that comes from the use of this complex theoretical model is that it will allow us to locate the break of the conceptual chain, providing a more accurate measure of the location of the semantic memory loss.

We will argue that all the patterns of semantic memory loss described in the literature can be properly accounted for by our model, which also makes it possible to put forward predictions on the correlation between different patterns of semantic memory loss, based upon location of the damage along the conceptual routes, and different types of dementia and brain damage.

References

Basso, A., E. Capitani & M. Laiacona (1988). Progressive language impairment without dementia: A case with isolated category specific semantic deficit. *Journal of Neurology, Neurosurgery, and Psychiatry* 51: 1202-1207.

Capitani, E., M. Laiacona, B. Mahon & A. Caramazza (2003). What are the facts of semantic category-specific deficits? A critical review of the clinical evidence. *Cognitive Neuropsychology* 20: 213-261.

Hillis, A.E. & A. Camarazza (1995). Category-specific naming and comprehension impairment: A double dissociation. *Brain* 114: 2081-2094.

Laws, K.R., R.L. Adlington, T.M. Gale, F.J. Moreno-Martínez & G. Sartori (2007). A meta-analytic review of category naming in Alzheimer's disease. *Neuropsychologia* 45: 2674-2682.

Martín Mingorance, L. (1998). *El modelo Lexemático-Funcional*. A. Marín (ed.). Granada: Publicaciones de la Universidad de Granada.

Peraita, H., C. Díaz & L. Anllo-Vento (2008). Processing of semantic relations in normal aging and Alzheimer's disease. *Archives of Clinical Neuropsychology. Elsevier.23*(1), 33-46.

Periñán C. & F. Arcas (2007). Cognitive modules of an NLP knowledge base for language understanding. *Procesamiento del Lenguaje Natural* 39: 197-204.

Silveri, C. & G. Gainotti (1988). Interaction between vision and language in category-specific semantic impairment. *Cognitive Neuropsychology* 5: 677-709.
Siri, S., E.A. Kensinger, S.F. Cappa & K.L. Hood, K.L. (2003). Questioning the Living/Non-living Dichotomy: Evidence from a Patient with an Unusual Semantic Dissociation. *Neuropsychology* 17.4: 630-645.
Warrington, E.K., & T. Shallice (1984). Category specific semantic impairments. *Brain* 107: 829-854.

Departamento de Filologías Extranjeras y sus Lingüísticas. UNED, 28040 Madrid, Spain.
bperez-cabello@flog.uned.es
iteomiro@flog.uned.es

Grammatical words as structural dominants in linguistic schematization of cognitive experience

Irina Tolmacheva

Center for Cognitive Research, Derzhavin Tambov State University, Tambov, Russia

The research addresses the problem of holistic description of structural and functional nature of linguistic cognition in its dependence on the individual-specific features of conceptual system configuration, those features determining the subjective nature of verbal forms of human cognitive and communicative activity. The core of the problem is that both individual's cognitive activity and the ways of presenting its products in language and discourse are defined by personal dominants in the structure of the individual's conceptual system.

The aim of the research is to explore the ways in which grammatical (or function) words contribute to the structural configuration of conceptual system and analyze the means of linguistic representation of grammatical meanings by function words.

We focus on the arrangement and rearrangement of meaning represented by grammatical words considering grammatical meanings to be a specific format of knowledge.

The research is based on the anthropocentric theory of language [Boldyrev 2015; Boldyrev 2017] and on the previous results of cognitive study of the function words category [Boldyrev, Tolmacheva 2014; Tolmacheva 2016, Tolmacheva 2017]. The cognitive and the anthropocentric approaches to the study of language and mind are regarded to be valuable tools for studying grammatical words from the perspective of schematizing cognitive experience through language. Within anthropocentric framework a grammatical word can be viewed as a means of conceptualization of the reality in the context of its structural organization in the way it is mentally cognized and interpreted in language.

The results expected involve identifying and describing those dominant constructs of linguistic cognition which determine the role of function words in both conceptualization and linguistic representation of the world.

References

Boldyrev, Nikolay & Tolmacheva, Irina. 2014. Categories of content and function words in language. In: Cognitive studies of language, 17, 369-375. (In Russ.)

Boldyrev, Nikolay. 2015. Anthropocentric nature of language in its functions, units, and categories. In: Voprosy kognitivnoy lingvistiki, 1, 5-12. (In Russ.)

Boldyrev, Nikolay. 2017. Anthropocentric theory of language. Theoretical and methodological foundations. In: Cognitive studies of language, 28, 20-81. (In Russ.)

Tolmacheva, Irina. 2016. Factors of secondary interpretation of language units used as function words. In: Cognitive studies of language, 27, 681-685. (In Russ.)

Tolmacheva, Irina. 2017. Anthropocentric nature of the function words category. In: Cognitive studies of language, 28, 263-277. (In Russ.)

Derzhavin Tambov State University, 33, Internatsionalnaya str., Tambov, 392000, Russia,
i_tolmacheva@mail.ru

ACKNOWLEDGEMENTS: THE RESEARCH IS FINANCIALLY SUPPORTED BY THE RUSSIAN SCIENCE FOUNDATION GRANT (PROJECT 18-18-00267) AT DERZHAVIN TAMBOV STATE UNIVERSITY.

Challenges for knowledge representation: Emergence in linguistic expressions and Internet Memes

Elke Diedrichsen
Institute of Technology, Blanchardstown, Dublin
e.diedric@googlemail.com

In modern approaches to linguistics, the relationship between signifier and signified is not believed to be something static, that is once and for all stored in the mental lexicon and shared by all speakers of a language. Rather, the concept of 'emergence' has entered the discussion of the way people create and understand linguistic items and utterances, and it seems to encompass all aspects of linguistic production and comprehension.

The idea that grammar with its categories is emergent is defended by Paul Hopper (2011), who claims that in order to understand grammatical categories, one has to look at corpus data and explain grammatical features in terms of actual usage, and not based on some abstract rules that only apply to made-up, written sentences. The nature of grammar as an emergent phenomenon also entails that the rules and categories are fleeting, and only observable in discourse.

As for the form of a linguistic sign, it has been stated that the form is "coined" (Feilke 1996, 1998) in language use, and that there is therefore no guarantee for form-meaning correlations to be regular. Philosophers of language (Wittgenstein 1960, Eco 1976) have maintained that the same holds for a sign's meaning. For the correct usage of a sign a speaker will need to know the usage conventions in a culture of speakers. The sign is therefore a cultural unit, and its meaning is defined by usage. It is also dynamically created in usage, as the semantics of constructions and lexical items are merely meaning potentials (Linell 2005, Croft and Cruse 2004).

These considerations suggest that speakers of a language acquire knowledge about words and constructions through exposure to a common culture. However, the concept of a 'culture' and the knowledge shared in it is dynamic as well. Kecskes (2008, 2010, 2012, 2014, Kecskes and Zhang 2009) maintains that the "Common Ground", which is the knowledge shared between speakers in an interaction, may but need not be shared in advance of the conversation. There is also "emergent common ground", which is knowledge that comes up as part of the interaction and is dynamically integrated by the interactants. As for culturally shared cognitive concepts that provide the basis for the behaviour in an interaction and the linguistic expressions used, Sharifian (2011, 2017) finds that these are emergent and negotiable as well.

The paper will discuss these dynamic approaches to communication as challenges to knowledge representation. The phenomenon of emergent forms and meanings will be exemplified by a modern form of Internet communication, the Internet Meme. Internet Memes are cultural phenomena, as they are instantiations of peer group knowledge. They reflect a peer group's main interest / trend / flavour of the month, and they incorporate sentiments shared by that peer group (Diedrichsen submitted).

The case of Internet Memes as communicative units shows that in the growing field of Internet communication, the form is not restricted to linguistic forms in spoken or written. The content merely has to be perceivable and distributable via the Internet, and any function or meaning may emerge through immediate mass distribution and interaction.

References:

Croft, W. and Cruse, A. 2004. *Cognitive Linguistics.* Cambridge: Cambridge University Press.
Diedrichsen, E. submitted. On the Interaction of Core and Emergent Common Ground in Internet Memes. To appear in *Internet Pragmatics,* special issue on the Pragmatics of Internet Memes.
Eco, U. 1976. *A theory of semiotics.* Bloomington, London: Indiana University Press.
Feilke, H. 1996. *Sprache als soziale Gestalt. Ausdruck, Prägung und die Ordnung der sprachlichen Typik,* Frankfurt am Main: Suhrkamp.

Feilke, H. 1998. Idiomatische Prägung. In Barz, I. & Öhlschläger, G. (Eds.): *Zwischen Grammatik und Lexikon.* Tübingen: Max Niemeyer, 69–80.

Hopper, P. J. 2011. Emergent Grammar and Temporality in Interactional Linguistics. In Auer, P. & Pfänder, S. (Eds.): *Constructions: Emerging and Emergent.* Berlin/New York: De Gruyter, 22–44.

Kecskes, I. 2008. Dueling Contexts: A Dynamic Model of Meaning. *Journal of Pragmatics* 40, 385–406.

Kecskés, I. 2010. The paradox of communication. In *Pragmatics and Society* 1:1, 50-73.

Kecskes, I. 2012. Sociopragmatics and Cross-Cultural and Intercultural Studies. In *The Cambridge Handbook of Pragmatics,* ed. by Keith Allan, and Kasia M. Jaszczolt, 599–616. Cambridge: Cambridge University Press.

Kecskes, I. 2014. *Intercultural Pragmatics.* New York: Oxford University Press.

Kecskes, I. & Zhang, F. 2009. Activating, seeking and creating Common Ground: A Socio-Cognitive Approach. *Pragmatics & Cognition* 17, 2, 331–355.

Linell, P. 2006. Towards a dialogical linguistics. In Lähteemäki, M., H. Dufva, S. Leppänen and P. Varis (eds.): *Proceedings of the XII International Bakhtin Conference Jyväskylä, Finland, 18–22 July, 2005.* University of Jyväskylä, Finland.

Sharifian, F. 2011. *Cultural Conceptualisations and Language: Theoretical framework and applications.* Amsterdam/Philadelphia: John Benjamins.

Sharifian, F. 2017. *Cultural Linguistics.* Amsterdam/Philadelphia: John Benjamins.

Wittgenstein, L. 1960. Philosophische Untersuchungen. In: Wittgenstein, L. *Tractatus logico-philosophicus (= Schriften 1),* Frankfurt am Main: Suhrkamp.

Referent tracking in narrative in three western desert dialects

Conor Pyle
Trinity College Dublin

This paper is a Role and Reference Grammar (Van Valin & LaPolla 1997) analysis of how referents are kept track of in text in Pitjantjatjara, Yankunytjatjara and Ngaanyatjarra (PYN). Role versus reference has two functions in syntax (N. Enfield p.c.), signalling the role of arguments with respect to the clause and with reference to what was said in previous clauses. Cross-linguistically, new referents are generally introduced in absolutive (S or O) roles (Du Bois 1987: 827, N. Enfield p.c.), because the A role is usually the topic and is referenced by a pronoun in the narrative, whereas the O argument is often ephemeral. In PYN, characters are introduced on first mention, thereafter pronoun clitics are used, being cognitively lighter than full pronouns: a zero 3rd person default clitic and ellipsis extend this trend, a null pronoun being a zero anaphor retaining salience from a previous clause. This leads to verb rich utterances, with verbs frequently in series. Thus an argument is backgrounded once it has been established in discourse, which is part of 'information flow' (Mithun 1999, Velázquez-Castillo 1995). PYN also has switch reference particles and sub-clauses which obviate the need for overt expression of syntactic subject.

We draw on 'Common Ground', which is mutual knowledge, beliefs and assumptions (Clark & Brennan 1991). As participants speak, they 'ground' what has been said in the conversation. There is a presupposition by the speaker of what is common ground (Stalnaker 2002). Thus a sentence may be appropriate only in a particular situation. Core common ground (including common sense and cultural knowledge) is distinguished from emergent common ground (Kecskes & Zhang 2009) which builds during a conversation. In small communities there is a high degree of local knowledge so no need to specify everything in conversation (Baker & Mushin 2008: 13-14), and cognate verbs imply the existence of an undergoer that does not need to be overtly expressed. Centering theory refers to the centre of attention in a conversation and this affects the form that referring expressions take (Thomason 2003, Walker, Joshi & Prince 1998: 1). Forward looking centres are discourse entities evoked by an utterance, while backward looking entities are similar to topics (Walker, Joshi & Prince 1998: 3). As conversation progresses the topics under discussion develop and change. Centering theory seeks to address anaphora resolution. There is rich information in a first utterance, but memory of utterances fades rapidly (Roberts 1998: 359-361) which means unless referents are constantly refreshed, they may need to be explicitly stated again.

PYN arguments thus do not need to be specified; though it leaves a sentence technically incomplete; and reference crucially depends on context. These may be accounted for by exophoric expressions deriving from the situation; endophoric ones referring to something already in the text or homophoric ones deriving their interpretation from cultural reference (D. Rose p.c.). RRG posits a completeness constraint whereby syntactic expression links to the semantic representation, and in this study we characterise how this can be accounted for in PYN.

References

Baker, B., & Mushin, I. (2008). Discourse and Grammar in Australian Languages. In I. Mushin & B. J. Baker (Eds.), Discourse and Grammar in Australian Languages (pp. 1–24). Amsterdam/Philadelphia: John Benjamins Publishing Company.

Clark, H. H., & Brennan, S. E. (1991). Grounding in communication. In L. B. Resnick, J. M. Levine, & S. D. Teasley (Eds.), Perspectives on socially shared cognition (pp. 127–149). Washington: American Psychological Association.

Du Bois, J. W. (1987). The Discourse Basis of Ergativity. Language, 63(4), 805–855.

Kecskes, I., & Zhang, F. (2009). Activating, seeking and creating common ground. A socio-cognitive approach. Pragmatics & Cognition, 17(2), 331–355. http://doi.org/10.1075/p

Mithun, M. (1999). The languages of native North America. Cambridge: Cambridge University Press.

Roberts, C. (1998). The place of centering in a general theory of anaphora resolution. In M. A. Walker, A. K. Joshi, & E. F. Prince (Eds.), Centering theory in discourse (pp. 359–399). Oxford: Clarendon Press.

Stalnaker, R. (2002). Common Ground. Linguistics and Philosophy, 25, 701–721. http://doi.org/10.1023/A:1020867916902

Thomason, L. G. (2003). The Proximate and Obviative in Meskwaki. The University of Texas at Austin.

Van Valin, Robert D. & LaPolla, Randy. J. 1997. *Syntax- Structure, Meaning and Function*. Cambridge: Cambridge University Press.

Velázquez-Castillo, M. (1995). Noun incorporation and object placement in discourse: The case of Guaraní. In P. Downing & M. Noonan (Eds.), Word Order in Discourse (pp. 555–579). Amsterdam/Philadelphia: John Benjamins.

Walker, M., Joshi, A., & Prince, E. (1998). Centering in naturally occurring discourse. In M. A. Walker, A. K. Joshi, & E. F. Prince (Eds.), Centering theory in discourse (pp. 1–28). Oxford: Clarendon Press.

Centre for Language and Communication Studies, Trinity College Dublin. pylec@tcd.ie

	KEYNOTE TALK Dr. Elizabeth Daly IBM

The role of previous discourse in detecting public textual cyberbullying

Aurelia Power

Antony Keane

Brian Nolan

Department of Informatics, Institute of Technology, Blanchardstown, Dublin, Ireland

Brian O'Neill

Graduate Research School, Dublin Institute of Technology, Dublin, Ireland

Previous work in the field of cyberbullying detection has focused solely on individual instances/posts taken in isolation, rather than part of the online conversation/dialogue. Consequently, the detection process typically considers only the information contained in the post itself, such as the presence of profane or violent words which may be indicative of cyberbullying. However, online discourse contains many instances that do not comply with grammatical standards, or they provide insufficient information (Crystal, 2011). For example, the instance *You clearly are* was labelled by annotators as cyberbullying in our dataset[1], despite the fact that its content suggests no cyberbullying, and it was only when we considered the previous post uttered by a different user - *I am not pathetic* - that we were able to identify one of the cyberbullying elements in the form of the offensive adjective *pathetic*. To address this limitation, we investigate here the role of previous instances/posts in identifying the missing cyberbullying elements, and we propose a framework that relies on the definition of cyberbullying that we posit elsewhere (Power et al, 2017; Power et al, in press) and on the information paradigm proposed by Prince (1981) who divides information into discourse-old and discourse-new. Specifically, the focus of the present paper is on how discourse-old information is used to infer some or all three necessary and sufficient cyberbullying elements: the personal marker, the dysphemistic element, and the link between them.

First, we analyse discourse-dependent instances of cyberbullying present in our dataset and propose a taxonomy of their underlying constructions as follows: (1) fully inferable constructions – where all three cyberbullying elements, the personal marker, the dysphemistic element, and the link between them, are not explicitly present, but can be inferred from previous discourse, (2) personal marker and cyberbullying link inferable constructions – where the dysphemistic element is explicitly present, but the personal marker and the link must be inferred from previous discourse, (3) dysphemistic element and cyberbullying link inferable constructions – where the personal marker is explicitly present, but the dysphemistic element and the cyberbullying link are entities inferable from previous discourse, and (4) dysphemistic element inferable constructions – where the personal marker and the link are explicitly present, but the dysphemistic element must be inferred from prior discourse. We then develop resolution rules to identify the personal marker, the dysphemistic element, and/or the cyberbullying link, in other words, to transform such instances into instances that contain them explicitly, and, therefore, into instances that can be subjected to the detection rules discussed elsewhere (Power et al, in press). We divide these resolution rules into separate sets that target the following: (1) polarity answers, (2) contradictory statements, (3) explicit ellipsis, (4) implicit affirmative answers, and (5) statements that use indefinite pronouns as placeholders for the dysphemistic element. Finally, we describe algorithms to implement these resolution rules, using several types of information: grammatical and syntactic information, such as part of speech and dependency relations among sentential constituents, as well as pragmatic information, such as the previous posts and the user names.

References

Allan, K. and Burridge, K. 2006. Forbidden Words: Taboo and Censoring of Language. Cambridge: Cambridge University Press.

Birner, B. 2013. Introduction to Pragmatics. Wiley-Blackwell Publishing.

Chatzakou, D., Kourtellis, N., Blackburn, J., De Cristofaro, E., Stringhini, G., and Vakali, A. 2017. "Mean Birds: Detecting Aggression and Bullying on Twitter." Cornell University Library: https://arxiv.org/abs/1702.06877.

[1]Our dataset originates from Ask.fm

Chen, Y., Zhou, Y., Zhu, S. and Xu, H. 2012. "Detecting Offensive Language in Social Media to Protect Adolescent Online Safety." Paper presented at the ASE/IEEE International Conference on Social Computing, 71 - 80. Washington, DC, September 3-5.

Crystal, D. 2011. Internet Linguistics: A Student Guide. Routledge, Taylor & Francis Group.

de Marneffe, M.C., and Manning, C. 2008b. "Stanford typed dependencies manual." https://nlp.stanford.edu/software/dependencies_manual.pdf.

Dinakar, K., Jones, B., Havasi, C., Lieberman, H., and Picard, R. 2012. "Common sense reasoning for detection, prevention, and mitigation of cyberbullying." ACM Transactions on Interactive Intelligent Systems, 2: 18:1-18:30. doi: 10.1145/2362394.2362400.

Dooley, J.J., Pyzalski, J., and Cross, D. 2009. "Cyberbullying versus face-to-face bullying – A theoretical and conceptual review." Journal of Psychology, 217: 182–188. doi: 10.1027/0044-3409.217.4.182.

Krifka, M. (2011). Questions. In Heusinger, Klaus von, Claudia Maienborn& Paul Portner (eds.), Semantics. An international handbook of Natural Language Meaning, Vol. 2. Berlin: Mouton de Gruyter, 1742-1785.

Langos, C. 2012. "Cyberbullying: The Challenge to Define." Cyberpsychology, Behavior, and Social Networks, 15(6): 285-289. doi: 10.1089/cyber.2011.0588.

Livingstone, S., Mascheroni, G., Ólafsson, K., and Haddon, L. with the networks of EU Kids Online and Net Children Go Mobile. 2014. "Children's online risks and opportunities: Comparative findings from EU Kids Online and Net Children Go Mobile". http://eprints.lse.ac.uk/60513/1/__lse.ac.uk_storage_LIBRARY_Secondary_libfile_shared_repository_Content_EU%20Kids%20Online_EU%20Kids%20Online-Children%27s%20online%20risks_2014.pdf.

Power, A., Keane, A., Nolan, B., and O'Neill, B. 2017. "A Lexical Database for Public Textual Cyberbullying Detection". Special issue of Revista de lenguas para fines específicos, entitled New Insights into Meaning Construction and Knowledge Representation.

Power, A., Keane, A., Nolan, B., and O'Neill, B. In press. "Detecting Discourse-Independent Negated Forms of Public Textual Cyberbullying". Journal of Computer-Assisted Linguistic Research.

Prince, E.F. 1981. Toward a taxonomy of given/new information. In P. Cole (ed.), Radical Pragmatics. Academic Press, 223-254.

Aurelia Power, Anthony Keane, Brian Nolan
Department of Informatics
Institute of Technology, Blanchardstown
Dublin 15
aurelia.power@itb.ie
anthony.keane@itb.ie
brian.nolan@itb.ie
Brian O'Neill
Graduate Research School
Dublin Institute of Technology
Dublin 7
brian.oneill@dit.ie

Implementing natural language understanding in an intelligent conversational agent

Gelmis S. Bartulis
Irene Murtagh
Institute of Technology Blanchardstown, Dublin, Ireland

This research work is concerned with the implementation of an intelligent conversational agent, also known as MIA (My Intelligent Agent). We use Natural Language Understanding (NLU), Text-to-Speech, Speech-to-Text and IBM Watson Conversation services, together with the IBM Bluemix platform in the development of this cognitively intelligent conversational agent. The objective of this Watson powered Android mobile application is to provide streamlined access to information and smartphone utilities based on the users interests and activities through the application of natural language understanding and cognitive technology.

The user converses with the agent using both speech input and text input, enhancing and streamlining the user experience. MIA has the ability to access the users social media content and provide an analysis of both the users sentiment and emotion by the application of semantic analysis. The agent provides ease of access to common daily tasks provided by an Android smartphone such as access to the news, accessing the users contacts and music collection, searching for nearby restaurants and shops based on the users current location and providing up to date weather information. The application can also provide a solution to the user based on the user intent. This solution can be derived from the users input using Natural Language Understanding (Cambria et al. 2014).

With the support of IBM Watson services and the IBM Bluemix platform we provide a cognitively intelligent conversational personal assistant that has the ability to converse with the user and provide solutions and answers to queries similar to a human personal assistant. This paper provides an overview of our approach in the implementation of this technology. We focus in particular on the NLU component within this application (Chopra et al. 2013).

References

Cambria, Erik and White, Bebo. 2014. Jumping NLP curves: A review of natural language processing research. Volume 9. IEEE Computational intelligence magazine, pp. 48–57.

Chopra, Abhimanyu, Prashar, Abhinav and Sain, Chandresh. 2013. Natural language processing. Volume 1. International Journal of Technology Enhancements and Emerging Engineering Research, pp. 131–134.

Room A15, Institute of Technology Blanchardstown, Dublin, Ireland.
B00080902@student.itb.ie
irene.murtagh@itb.ie

An experimental review on methods for Word Sense Disambiguation on Natural Language Processing

Fredy Núñez Torres

Pontificia Universidad Católica de Chile, Santiago, Chile

The following proposal presents a review and testing for the most relevant Word Sense Disambiguation (WSD) methods used nowadays on Natural Language Processing (NLP). This approach considers the development of experiments applied to a chilean spanish corpus that was designed based on the semantic representations available on the lexico-conceptual knowledge base FunGramKB (Periñán-Pascual and Arcas Tunez, 2004; Periñán-Pascual and Mairal-Usón, 2009). The main goal is to present and compare computational procedures for automatic WSD, such as machine learning (Pedersen, 2000; Zheng-tao et al., 2009); path-based metrics and overlapping glosses (Lesk, 1986; Resnik, 1995; Patwardhan et al., 2003); and multinomial logistic regression.

A semi-automatic selection of potentially polysemous lexical units (nouns) was carried out. In total, 120 instances (sentence context) were selected for 3 lexical units extracted from the written mass media corpus belonging to CODIDACH: Corpus Dinámico del Español de Chile (Dynamic Corpus of Chilean Spanish), development by Sadowsky (2006). Along with this, the selected lexical units were linked with specific concepts of the #ENTITY subontology of FunGramKB, as shown in the following table:

Lexical unit	**Concept (senses)**	**Description in FGKB**
Head (cabeza)	+HEAD_00	The upper or front part of the body in animals; contains the face and brains; "he stuck his head out the window"
	+INTELLIGENCE_00	Your ability to think, feel, and imagine things.
	+CHIEF_00	A person who is in charge; "the head of the whole operation"
	+LEADER_00	A person who rules or guides or inspires others.
Face (cara)	+FACE_00	The front of the head from the forehead to the chin and ear to ear; "he washed his face"; "I wish I had seen the look on his face when he got the news"
	+SIDE_00	A surface forming part of the outside of an object; "he examined all sides of the crystal"; "dew dripped from the face of the leaf"
Letter (carta)	+LETTER_00	A written message addressed to a person or organization.
	+CARD_00	A small piece of thick stiff paper with numbers or pictures on them, used to play a particular game.
	$MENU_00	A list of dishes available at a restaurant.

Table 1. *Polysemous lexical units and their corresponding concepts in FunGramKB*

The assembly and execution of all the experiments has been carried out using Data Mining Encountered, DAMIEN (Periñán-Pascual, 2017), a computer environment to support linguistic research. DAMIEN manages to integrate in the same work environment the different tools and techniques that can be applied in the analysis of linguistic corpora. These techniques come from different disciplines, such as Corpus Linguistics (e.g. frequency lists; XML processing and XSL; database administration and SQL; regular expressions; etc.), Statistics (e.g.

descriptive and inferential statistics; graphic representation of data; etc.), Natural Language Processing (e.g. extraction of n-grams; derivation; morphological and syntactic analysis; POS tags; etc.), and Text Mining (e.g. classification and clustering).

References

Lesk, Michael. 1986. Automatic sense disambiguation using machine readable dictionaries: How to tell a pine cone from a ice cream cone. Proceedings of SIGDOC '86.

Patwardhan, Siddhart; Banerjee, Satanjeev & Pedersen, Ted. 2003. Using Measures of Semantic Relatedness for Word Sense Disambiguation. Proceedings of the Fourth International Conference on Intelligent Text Processing and Computational Linguistics, pp. 241–257.

Pedersen, Ted. 2000. A Simple Approach to Building Ensembles of Naive Bayesian Classifiers for Word Sense Disambiguation. Proceedings of the North American Chapter of the Association for Computational Linguistics (NAACL), pp. 63-69.

Periñán-Pascual, Carlos. 2017. Bridging the gap within text-data analytics: a computer environment for data analysis in linguistic research. Revista de Lenguas para Fines Específicos 23.2 (2017), pp. 111-132.

Periñán-Pascual, Carlos & Arcas Túnez, Francisco. 2004. Meaning postulates in a lexico-conceptual knowledge base. Proceedings of the 15th International Workshop on Databases and Expert Systems Applications California, Los Alamitos: IEEE, 38-42.

Periñán-Pascual, Carlos & Marial-Usón, Ricardo. 2009. Bringing Role and Reference Grammar to natural language understanding. Procesamiento del Lenguaje Natural, 43, 265-273.

Resnik, Phillip. 1995. Using information content to evaluate semantic similarity. Proceedings of the 14th International Joint Conference on Artificial Intelligence, Montreal, Canada.

Sadowsky, Scott. 2006. Corpus Dinámico del Castellano de Chile (Codicach). Electronic database. http://sadowsky.cl/codicach.html

Zheng-tao, Yu; Bin, Den; Bo, Hou; Lu, Han & Jian-yi, Guo. 2009. Word Sense Disambiguation Based on Bayes Model and Information Gain. Proceedings of the International Journal of Advanced Science and Technology, Vol.3, February.

Language Sciences Department. Pontificia Universidad Católica de Chile. Santiago, Chile. *frnunez@uc.cl*

Discovering hazards via twitter for emergency management: a knowledge-based approach

Carlos Periñán-Pascual

Applied Linguistics Department, Universitat Politècnica de València, Spain

Hazards and disasters give rise to three main types of costs: (a) human cost, since they cause significant suffering and loss of lives, (b) economic cost, since they may result in damage and loss of property, and (c) environmental cost, since they can destroy natural habitats or release pollutants. Processing micro-texts from Twitter and other social media has become very valuable for the real-time detection of events that can affect our lives. Indeed, the automatic detection of such events can be really useful not only for citizens but also for emergency responders. The use of social sensors for the development of emergency response systems has become a relevant research topic over the last decade. The implementation of these systems usually has two characteristics in common. On the one hand, most of these systems were aimed at detecting a single or a few events, e.g. earthquakes (Sakaki et al. 2010, 2013; Liu et al. 2012), grassfires and floods (Vieweg et al. 2010) or swine flu (Signorini et al. 2011), among others. On the other hand, these systems employed a supervised machine-learning method of tweet classification (e.g. Naïve Bayes or SVM).

This research evaluates the performance of a knowledge-based system that exploits Twitter users as social sensors for the detection of multiple environmental hazards. In particular, we focus on the natural language processing module of the system, providing an account of the procedure to detect hazards from micro-texts: pre-processing tweets, discovering relevant features, determining topic and sentiment, and detecting the problem. Moreover, we explore the knowledge base developed for the system, since the degree of success of a symbolic approach is closely dependent on the quality and coverage of the lexical resources involved in the system. Finally, we draw attention to the advantages of our model over the machine learning approach. In this regard, our system is able not only to measure more effectively how reliable we can feel that a given tweet deals with some environmental hazard but also to set alert thresholds from which the severity of the problem could be rated.

References

Liu, S.B., Bouchard, B., Bowden, D.C., Guy, M. and Earle, P. 2012. USGS tweet earthquake dispatch (@USGSted): using twitter for earthquake detection and characterization, AGU fall meeting abstracts 1.

Sakaki, T., Okazaki, M. and Matsuo, Y. 2010. Earthquake shakes twitter users: real-time event detection by social sensors, Proceedings of the 19th international conference on World Wide Web ACM.

Sakaki, T., Okazaki, M. and Matsuo, Y. 2013. Tweet analysis for real-time event detection and earthquake reporting system development, IEEE Transactions on Knowledge and Data Engineering 25-4, 919-931.

Signorini, A., Segre, A.M. and Polgreen, P.M. 2011. The Use of Twitter to Track Levels of Disease Activity and Public Concern in the U.S. during the Influenza A H1N1 Pandemic. PLoS ONE 6 (5).

Vieweg, S., Hughes, A.L., Starbird, K. and Palen, L. 2010. Microblogging during two natural hazards events: what twitter may contribute to situational awareness. Proceedings of the SIGCHI conference on human factors in computing systems. ACM, 1079-1088.

Applied Linguistics Department. Universitat Politècnica de València, Spain. jopepas3@upv.es

ABSTRACTS – 5TH JULY, THURSDAY

Entrenchment of triconstituent English noun compounds

Elisabeth Huber
Ludwig Maximilian University of Munich, Germany

Why does *football* combine productively with further nouns to form more complex expressions like *football game*, whereas seemingly comparable compounds like *keyword* only seldom expand to more complex sequences? This project explores why some two-noun compounds are more readily available as a schema for forming triconstituent constructions than others. I hypothesize that the productivity of a schema for the formation of triconstituent sequences (e.g., '*football* + N') depends on the degree of entrenchment of the embedded two-noun compound (e.g. *football*), assuming that only strongly entrenched compounds are productive in forming more complex constructions.

Entrenchment in this project is measured through statistical measures (cf. Stefanowitsch & Flach 2016) based on usage frequencies extracted from the *Corpus of Contemporary American English*. The results of a preliminary study indicate a correlation between the entrenchment of two-noun compounds and their productivity as a schema. This suggests that the more entrenched a two-noun compound is, the more available it is as a schema to form more complex constructions. By contrast, two-noun compounds with low degrees of entrenchment will neither form many triconstituent constructions, nor will they be available for the production of new types.

Based on this result, a qualitative distinction between two fundamentally different kinds of triconstituent constructions becomes necessary: For the first type speakers have a schema available that consists of a strongly embedded two-noun compound and a slot for the third noun that is variable to a stronger or lesser degree (e.g. *football* + N). The second type of triconstituent construction is not based on such schema but probably acquired as ready-made term (e.g. *body mass index*). Its mental representation is thus not schematic but the construction is explicitly filled on the lexical side.

This insight will open the ground for further discussion on the wordhood of triconstituent sequences. Can they reach the status of one cognitive unit despite their complexity? The experimental literature strongly tends to equate strong entrenchment (i.e. high processing speed) with holistic chunking (cf. Blumenthal-Dramé 2016). This project will argue that there is a more sensible approach to the sequences under investigation in this project than through the concept of chunking, as they might not necessarily be stored as one unit. With the help of predictability measurements, I will make use of the theory of entrenchment (Schmid 2007) to explain how the syntagmatic associations between a compounds' constituents grow strong enough that the activation of an embedded compound can strongly trigger a third noun that is repeatedly combined with it. This cognitive approach will bring new input into the discussion of whether or not noun sequences can be considered compounds, i.e. one word (cf. Bauer 1988), as compounding has traditionally been notoriously hard to pin down between the interfaces of lexis, morphology and syntax.

References

Bauer, L. (1988). When is a sequence of two nouns a compound in English? *English Language and Linguistics, 2*(1), pp. 65-86.

Blumenthal-Dramé, A. (2016). Entrenchment from a psycholinguistic and neurolinguistic perspective. In H.-J.Schmid (ed.), *Entrenchment and the psychology of language learning: how we reorganize and adapt linguistic knowledge,* pp. 101-127. Berlin: Walter de Gruyter.

Schmid, H.-J. (2007). Entrenchment, salience and basic levels. In D. Geeraerts & H. Cuyckens (eds.), *The Oxford Handbook of Cognitive Linguistics*, Oxford: Oxford University Press, pp. 117-138.

Schmid, H.-J. (2016). *Entrenchment and the psychology of language learning. How we reorganize and adapt linguistic knowledge*. Berlin: Walter de Gruyter.

Stefanowitsch, A., & Flach, S. (2016). A corpus-based perspective on entrenchment. In H.-J. Schmid (ed.), *Entrenchment and the psychology of language learning: how we reorganize and adapt linguistic knowledge,* pp. 101-127. Berlin: Walter de Gruyter.

Department of English Linguistics. Ludwig Maximilian University of Munich, Munich D-80539.
huber.elisabeth@lmu.de

Parallels and contrasts between the approved adjectives in the ASD-STE dictionary and the adjectival concepts in FunGramKB Core Ontology

Ángela Alameda Hernández
Ángel Felices Lago
Department of English and German, University of Granada, Granada, Spain

In a previous contribution to the MKR 2017 Conference we offered the results of comparing and matching basic and terminal verbal concepts under the subontoloy #EVENT in FunGramKB Core Ontology (and their corresponding lexical units in the English Lexion) with Words [Approved Verbs] (and their corresponding synonyms) in the ASD-STE dictionary. As a consequence, we intended to determine whether the way in which this controlled language had been designed could draw similarities with the way in which the conceptual information of FunGramKB (Knowledge Base) had been built. The outcome was that more than half of the Approved Verbs in the STE Dictionary (58%) were also represented in FunGramKB, either as concepts or as lexical units associated to other concepts. This was a surprisingly high percentage of matching, due to the fact that the ASD-STE Dictionary had been basically conceived as a repository to help aircraft maintenance workers and, inevitably, a high number of the units included in the dictionary should be connected directly or indirectly with the characteristics of the language and the lexical resources employed by these aircraft technicians.

At present, we intend to observe whether this tendency for matching only affects verbal concepts or it is also replicated in other FunGramKB subontologies under the Core Ontology. To provide evidence based on authentic material, we have selected the list of 226 Approved Adjectives in the ASD-STE dictionary: a collection of units (the same as verbs, nouns, adverbs, etc.) complying with the ASD-STE lexical and syntactic restrictions. These adjectives are used as a representative sample to be compared with 329 adjectival concepts stored in the FunGramKB #QUALITY subontology (as basic or terminal concepts). The level of compatibility between both repositories may offer four possibilities of conceptual and/or lexical matching at varying degrees: i) direct matching, ii) indirect matching, iii) no matching, or iv) missing. The quantitative results of this analysis may also prove that a significant percentage of Adjectival Words in the ASD-STE dictionary are directly or indirectly represented in FunGramKB, either as concepts or as lexical units associated with other concepts, even if the level of matching is lower than in the case of the verbal concepts. In addition to that, the analytical comparison between adjectival concepts and units in both repositories may also demonstrate how the ASD-STE Dictionary can be of help to improve and extend the FunGramKB Core Ontology, which is still under construction.

References

ASD Simplified Technical English Specification ASD-STE 100 TM: International specification for the preparation of maintenance documentation in a controlled language. Issue 7. January 2017. Brussels: ASD.

Felices, Ángel & Ureña, Pedro. 2012. Fundamentos metodológicos de la creación subontológica en FunGramKB. Onomázein, 26 (2), pp. 49-67.

Felices, Ángel & Alameda, Ángela. 2017. The process of building the upper-level hierarchy for the aircraft structure ontology to be integrated in FunGramKB. Revista de Lenguas para Fines Específicos 23 (2), pp. 86-110.

Felices, Ángel & Alameda, Ángela. Forthcoming. Parallels and contrasts between the ASD-STE Dictionary and the Ontology in FunGramKB. Onomázein.

Kuhn, Tobias. 2014. A Survey and Classification of Controlled Natural Languages. Computational Linguistics, 40(1), pp. 121-170.

Periñán, Carlos. & Arcas, Francisco. 2004. Meaning postulates in a lexico-conceptual knowledge base. In 15th International Workshop on Databases and Expert Systems Applications. IEEE, Los Alamitos (California), pp. 38-42.

Periñán, Carlos & Arcas, Francisco. 2010a. Ontological commitments in FunGramKB. Procesamiento del Lenguaje Natural, 44, pp. 27-34.

Periñán, Carlos & Arcas, Francisco. 2010b. The architecture of FunGramKB. In Proceedings of the 7th International Conference on Language Resources and Evaluation. European Language Resources Association, pp. 2667-2674.

Periñán, Carlos & Mairal, Ricardo. 2010. La Gramática de COREL: un lenguaje de representación conceptual. Onomázein. 21 (1), pp. 11-45.

Reiss, M., Moal, M., Barnard, Y., Ramu, J. Ph., & Froger, A. 2006. Using ontologies to conceptualize the aeronautical domain. In Proceedings of the International Conference on Human-Computer Interaction in Aeronautics (HCI-AERO 2006), Seattle, WA, Toulouse: Cépaduès-Editions, pp. 56-63.

Department of English and German, University of Granada, Campus de Cartuja s/n, 18071 Granada, Spain.
aalameda@ugr.es

OVER in radiotelephony communications

Maria del Mar Robisco Martin

Department of Linguistics applied to Science and Technology, Technical University of Madrid

For over sixty years aviation radiotelephony has been based on a standard phraseology designed to achieve the utmost clarity and brevity and to minimise failures in air-ground communication. It consists of codified and limited dialogues between air traffic controllers and flight crew members. In 1997, the *International Civil Aviation Organization (ICAO)* created the *Proficiency Requirements in the Common English Study Group* and in March 2003, it was decided that 1) English should be the universal medium for radiotelephony communications. 2) All pilots and controllers should pass an English language exam to achieve an operating level[2]. 3) All pilots and controllers should make a global use and a correct application of the phraseology in these interactions. 4) It would be necessary to carry out studies to analyse the English language in these communications and to create teaching resources.

Thus, following the *ICAO's Requirements,* this paper focuses on the language employed in radiotelephony communication, in particular, on the preposition *over*. This study is significant because a) the polysemy can affect the interpretation of a sentence and create misunderstandings b) prepositions are amongst the most polysemous words in English c) the semantic network associated with any preposition in one language rarely overlaps with the meanings of any single linguistic form in another language (Taylor 1989, 112) and d) *over* is perhaps the most polysemous of the English prepositions (Taylor 1989, 113). The aim of the paper is to show the multiplicity and fuzziness of meaning in natural language, in contrast with the simplified view suggested by the standard phraseology. The purpose is twofold: First, to demonstrate that *over* appears with more meanings than with the Primary Sense which is the only meaning included in the radiotelephony phraseology and, secondly, to systematize the senses of *over* in this context, which constitute a complex network of related meanings. Tyler and Evans (2003:90) claim that "the distinct senses associated with spatial particles arise through experiential correlation (the same correlation that motivates the metaphor) and strengthening of implicatures that arise in the course of sentence interpretation".

An electronic database consisting of 69 cockpit voice recordings[3] was used to produce an electronic corpus of 3226 items. The recordings are taken from fatal aviation accidents which occurred between 1962 and 2002. Although it is limited, a clear pattern seems to have emerged. The findings suggest that the Primary Sense is the only meaning of *over* proposed by the phraseology whereas an examination of aircraft communications shows that o*ver* is used in a range of other senses (the Primary Sense, the Covering Sense, the On-the-other-side Sense, the Transfer Sense, the Completion Sense, the More Sense, the Control Sense, the Examination Sense and the Repetition Sense) which create a semantic network.

References

Brugman, C. (1981). "Story of Over". MA thesis, University of California, Berkely.

Brugman, C. (1988). The Story of Over: Polysemy Semantics and the Structure of the Lexicon. New York Garland Publishing, Inc

Cuenca, M.J. & Hilferty, J. (1999). Introducción a la lingüística Cognitiva. Ariel Lingüística.

Dirven, R. (1993). "Dividing up Physical and Mental Space into Conceptual Categories by Means of English Prepositions". In Zelinsky-Wibbelt, Cornelia (ed.), 73-97.

Hinton, H, Paunero, N., Martinez, J., Velasco, A. & Gomez, F. (2002). Manual de Comunicaciones Aeronáuticas. Cockpitstudio.

Johnson, M. (1987). The Body in the Mind: The Bodily Basis of Meaning, Reason and Imagination. Chicago: University Chicago Press.

Lakoff, G. (1979). "The Contemporary Theory of Metaphor". In Ortony (ed.), 202-251.

Lakoff, G. (1987). Women, fire and dangerous things. Chicago: University Chicago Press

Lakoff, G. & Johnson, M. (1980). Metaphors We Live By. Chicago: University Chicago Press.

Lakoff, G. & Johnson, M. (1999). Philosophy in the Flesh: the Emboided Mind and its Challenge to Western Thought. Chicago: University Chicago Press.

Langacker, R.W. (1987). "Foundations of Cognitive Grammar". *Vol. I: Theoretical Prerequisites.* Stanford:

[2] ICAO operational level: An operating level equivalent to level four, in a ranking from one to six

[3] www.planecrashinfo.com. They belong to an aviation accident database which includes all civil aviation accidents of scheduled and non-scheduled passenger airliners worldwide, which resulted in at least one fatality.

Stanford University Press.
McCarthy, M. (1991). Discourse Analysis for Language Teachers. Cambridge: Cambridge University Press.
Rosch, E. (1975). "Cognitive Representations of Semantic Categories". *Journal of Experimental Psychology*: *General* 104, 193-233.
Taylor, J.R. (1989). Linguistic Categorization. New York: Oxford University Press.
Tyler, A. & Evans, V. (2003)."Reconsidering Prepositional Polysemy Networks: The Case of *Over*". In B. Nerlich, Z. Todd, V. Herman, & D. Clarke (Eds), *Polysemy: Flexible patterns of meanings in mind and language* (pp95-160). Reprinted from *Language*, 77, 4, 724-765. Berlin: Mouton de Gruyter.
Tyler, A. & Evans, V. (2004)."Applying Cognitive Linguistics to Pedagogical Grammar: The Case of *Over*". In M. Achard and S. Niemeier. *Cognitive Linguistics, Second Language Acquisition, and Foreign Language Teaching*, 257-280. Berlin: Mouton de Gruyter.
Wittgenstein, L. (1978). Philosophical Investigations. Oxford: Basil Blackwell

ETSIAE, Plaza de Cardenal Cisneros 3, 28040 Madrid. mariadelmar.robisco@upm.es

Mechanisms of metaphtonymy formation (based on English verbs with semantics “to separate”)

Svetlana Kiseleva
Humanitarian Department, Saint Petersburg State University of Economics, Saint Petersburg, Russia
Nella Trofimova
Department of Foreign Languages, National Research University Higher School of Economics, Saint Petersburg, Russia
Irina Rubert
Humanitarian Department, Saint Petersburg State University of Economics, Saint Petersburg, Russia

The paper analyzes the interaction of metaphor and metonymy, known as metaphtonymy, and its functioning in the context on the basis of verbs with semantics “to separate”. It discusses the main models of metaphtonimic projection: metaphor and metonymy; metonymy–metaphor–metonymy; metaphor based on metonymy (partially or fully); metonymy based on metaphors. The relevance of this study lies in the lack of study of cognitive values from the standpoint of metaphor and metonymy interaction in conditions of intersection of verbs close in meaning with semantics “to separate”. The novelty of this work lies, firstly, in the consideration of the mechanism of formation of the basic cognitive schemas of metaphtonymic meanings, in how the phrase can acquire a new or additional meaning depending on the location of words in the context, and secondly, it is the study of the mechanism of metaphtonymy formation in conditions of intersection of close verbs with the semantics “to separate”.

Metaphors and metonymies are effective means of conceptualizing new elements of the modern worldview, since as concepts become more complex, the mechanisms of naming the surrounding reality become more complex too. Metaphtonymy is an example of such more complex structures. The basis of metaptonymy (the term is proposed by L. Goossens (1990)) is based on the principles of integration processes of metaphorical and metonymic blending. Such a complex unit can combine the properties of both metaphors and metonyms. More recent studies have provided more refined and systematic patterns of interaction between metaphor and metonymy (cf. Ruiz de Mendoza and Galera-Masegosa, 2011). However, our corpus of analysis suggests that further developments are needed in order to fully account for the complexities of verb with semantics of separation interpretation.

Following J. Lakoff, L. Goossens, metaphor is considered as the projection of elements of different conceptual domains: the source domain and the target domain, metonymy is understood as a projection of adjacent elements of one conceptual domain [Lakoff, 1987; Gossens, 2002]. A cognitive approach to analysis of metaphor and metonymy can be considered as conceptual interaction in the complex and reach to metaphtonimic modeling. Also this approach reveals the interaction of metaphors and metonymy as a complex mechanism of the formation of meanings, as realized in context. The results of this study can contribute to the theory of metaphor, metonymy, secondary language nomination.

References

Goossens, Louis. 2002. Metaphtonymy: The interaction of metaphor and metonymy in expressions for linguistic action // Dirven R, Pörings R. Metaphor and Metonymy in Comparison and Contrast. N.Y., Berin: Walter de Gruyter, pp. 349-378.

Lakoff, John. 1987. Women, fire and dangerous things: What categories reveal about the mind. Chicago; London.Ruiz de Mendoza, Francisco Jose Ruiz; Galera-Masegosa, Alicia. 2011. “Going beyond metaphtonymy: Metaphoric and metonymic complexes in phrasal verb interpretation”. Language Value, 3 (1), 1-29. Jaume I University ePress: Castelló, Spain.

Humanitarian Department, Saint Petersburg State University of Economics, 195023, St. Petersburg, Moskatelny per. 4, Russia, svkiseljeva@bk.ru

Department of Foreign Languages, National Research University Higher School of Economics, 198099, St. Petersburg, ul. Promyshlennaya, 17, Russia, nelart@mail.ru

Sampling techniques to overcome class imbalance in a cyber bullying context

David Colton
IBM, Dublin, Ireland
Markus Hofmann
Department of Informatics, Institute of Technology Blanchardstown, Dublin, Ireland

The majority of datasets suffer from class imbalance where samples of a dominant class, significantly outnumber the samples available for the minority class that is to be detected. Prediction and classification machine learning models work best when there are roughly equal numbers of each class type. This paper explores sampling techniques that can be used to overcome this class imbalance problem in a cyber bullying context.

A newly classified cyberbullying dataset, including detailed descriptions of the criteria used in its classification, was generated to examine the feasibility of using text mining techniques, to automate the detection of cyberbullying text. When the dataset shows a significant class imbalance between the positive, cyberbullying, sample and the negative, not cyberbullying, samples.

In this paper, we will investigate if over sampling the minority positive class or under sampling the majority negative class affects the performance of a prediction model. A compromise solution where the positive class is partially over sampled, and the negative class is partially under sampled is also examined. More advanced methods for handling class imbalance such as cost based learners or the Synthetic Minority Over-sampling Technique (SMOTE) were not considered at this time. Although not strictly a class imbalance solution, sampling using the most frequently observed features is also explored.

Motivating the computational phonological parameters of an Irish Sign Language avatar

Irene Murtagh
Institute of Technology Blanchardstown, Dublin, Ireland

This paper provides an account of the computational phonological parameters of an Irish Sign Language (ISL) avatar. We provide a motivation of the phonological-morphological interface in ISL. This work is part of research work in progress in the development of a linguistically motivated computational framework for ISL. We use Role and Reference Grammar (RRG) (Van Valin and LaPolla 1997) as the theoretical framework of this study. Using RRG provides significant theoretical and technical challenges within both RRG and software (Van Valin 2005).

Prior to preparing a linguistically motivated computational definition of lexicon entries that are sufficient to represent ISL within the RRG lexicon we must first define ISL phonological parameters in computational terms. Due to the visual gestural nature of ISL, and the fact that ISL has no written or aural form, in order to communicate an ISL utterance in computational terms we must implement the use of a humanoid avatar capable of movement within three-dimensional (3D) space. In providing a definition of a linguistically motivated computational model for ISL we must be able to refer to the various articulators (hands, fingers, eyes, eyebrows etc.), as these are what we use to articulate various phonemes, morphemes and lexemes of an utterance (Murtagh 2011a, 2011b).

We propose a new level of lexical representation (Pustejovsky 1995), which describes the essential (computational) phonological parameters of an object as defined by the lexical item. Our proposed new level of lexical meaning: articulatory structure level, caters specifically for the computational linguistic phenomena consistent with signed languages, in particular to this research ISL, enabling us to adequately represent ISL within the RRG lexicon.

References

Murtagh, I. 2011a. Developing a Linguistically Motivated Avatar for ISL Visualisation. In: Workshop on Sign Language Translation and Avatar Technology. [online] Dundee: University of Dundee, Scotland. Available at: http://vhg.cmp.uea.ac.uk/demo/SLTAT2011Dundee/7.pdf [Accessed 17.03.18].

Murtagh, I. 2011b. Building an Irish Sign Language Conversational Avatar: Linguistic and Human Interface Challenges. Conference on Irish Human Computer Interaction. Cork: Cork Institute of Technology, Ireland.

Pustejovsky, J. 1995. The generative lexicon. Cambridge: MIT Press.

Van Valin, R. 2005. Exploring the Syntax-Semantics Interface. Cambridge: Cambridge University Press.

Van Valin, R. and La Polla, R. 1997. Syntax: Structure, Meaning and Function. Cambridge: Cambridge University Press.

Room A15 Department of Informatics, Institute of Technology Blanchardstown, Blanchardstown Road North, Dublin 15, Ireland. murtagh.irene@gmail.com

Motivating a linguistically orientated model for a conversational software agent

Dr Kulvinder Panesar

School of Art, Design and Computer Science, York St John University, York, UK

This paper proposes a linguistically orientated model of a conversational software agent (CSA) (Panesar, 2017) framework sensitive to natural language processing (NLP) concepts and the levels of adequacy of a functional linguistic theory. We discuss the relationship between natural language processing and knowledge representation (KR), and connect this with the goals of a linguistic theory (Van Valin and LaPolla, 1997), in particular Role and Reference Grammar (RRG) (Van Valin Jr, 2005a). We discuss the advantages of RRG and fitness-for-purpose for computational implementation and its level of computational adequacy (Nolan, 2004). We propose a design of a computational model of the linking algorithm that utilises a speech act construction as a grammatical object (Nolan, 2014a, Nolan, 2014b) and the sub-model of belief-desire and intentions (BDI) (Rao and Georgeff, 1995). This model has been successfully implemented in software (Panesar, 2017, Pokahr et al., 2014), using conceptual graphs, and resource description framework (RDF), and we highlight some implementation issues that arose at the interface between language and knowledge representation.

Keywords: conversational software agents, natural language processing, speech act construction, knowledge representation, belief-desire and intentions, functional linguistics

References

NOLAN, B. 2004. First steps toward a computational RRG. *RRG2004 Book of Proceedings, ed. by Brian Nolan***,** 196-223.

NOLAN, B. 2014a. Constructions as grammatical objects : A case study of prepositional ditransitive construction in Modern Irish. *In:* NOLAN, B. & DIEDRICHSEN, E. (eds.) *Linking Constructions into Functional Linguistics: The role of constructions in grammar.* Amsterdam/Philadelphia: John Benjamins Publishing Company.

NOLAN, B. 2014b. Extending a lexicalist functional grammar through speech acts, constructions and conversational software agents. *In:* NOLAN, B. & PERIÑÁN., C. (eds.) *Language Processing and Grammars: The role of functionally oriented computational models [Studies in language Companion Series 150].* Amsterdam and New York: John Benjamins Publishing Company.

PANESAR, K. 2017. *A linguistically centred text-based conversational software agent.* Unpublished PhD thesis, Leeds Beckett University.

POKAHR, A., BRAUBACH, L., HAUBECK, C. & LADIGES, J. 2014. Programming BDI agents with pure Java. *Multiagent System Technologies.* Springer.

VAN VALIN JR, R. D. 2005a. *Exploring the syntax-semantics interface*, CUP.

School of Art, Design and Computer Science, York St John University, k.panesar@yorksj.ac.uk

Speaker's focus of interest as a basis of a text semantic model

Irina Ivanova-Mitsevich
Independent Researcher, Warsaw, Poland

As it was proved by M. Bachtin all texts are elements of dialogues (Bachtin 1986). From that point of view a text might be treated as a communicative step created by a number of speech acts (Austin 1962). Thus, a text should have pragmatic unity and semantic coherence. Pragmatic unity is coordination of the illocutionary forces of the speech acts of the text determined by the general illocution of the communicative step. Semantic coherence is coordination of different propositional content of speech acts (Vanderveken 2009; Searle 1969) or in other words different fragments of the picture of the world that constitute the subject matter of the text. The problem is how these two lines of textual unity interact, in fact how the semantic coherence complies with the pragmatic aim of the text.

Traditionally, it is the lexical components of texts that are viewed upon as the link between pragmatic and semantic spheres of the text. Still the text is not only a collection of words but it is also a structural unity, thus the connection between the two spheres mentioned above should be somehow reflected in the structure of the text elements.

The minimal text unit that reflects a fragment of the picture of the world (a component of semantic coherence) and at the same time presents a speech act (the smallest component of pragmatic unity) is a sentence. Thus, in order to find the structural connection between the semantic and the pragmatic properties of a text it is necessary to reveal such an element of the sentence structure that is determined by the illocutionary force and in its turn determines the semantic and formal structure of the sentence.

The semantic structure of the English sentence might be presented as a frame consisting of the nominal and the predicate slots. The pragmatically loaded slot firstly, should freely be filled in by the components of the picture of the world because the choice of elements for this slot depends only upon the speaker's will, secondly, filling in of this slot should influence the distribution of other nominal elements within the frame.

Analysis of English sentences permits us to state that the slot situated just after the predicate can be filled in with practically any name of the components of the picture of the world and the choice of the name to fill in this slot restricts distribution of other names among other slots. The element of the picture of the world that appears in this slot might be named "the focus of speaker's interest".

Each sentence possesses a focus of interest. Since sentences are structural components of the text their foci of interest produce a net that on the one hand is the basis of semantic coherence because it presents most important element of the picture of the world reflected in the text and on the other hand it is part of the pragmatic unity because the choice of each focus of interest within a net is pragmatically determined.

References

Austin, J.L. 1962. How to Do Things with Words. Oxford: Oxford University Press.
Bachtin, Michał 1986. Estetyka twórczości słownej. Warszawa.
Searle, John R. 1969. Speech Acts: An Essay in the Philosophy of Language. New York: Cambridge University Press.
Vanderveken, Daniel. 2009. Meaning and Speech Acts: Volume 1, Principles of Language. Cambridge: Cambridge University Press.

ul. XXI Wieku nr 11B m. 2, 05-509 Józefosław, woj.mazowieckie Polska (Poland). irinka.ivanova@tut.by

From walled off Europe to walled in identity

Natalia Iuzefovich
Pacific National University
iuzefovichnata@gmail.com

The paper presents some results of an ongoing research project on the construction and reconstruction of identity, both national and individual.

The research is interdisciplinary and as such it combines the methods of socio-cognitive linguistics, psycholinguistics and corpus linguistics; both quantitative and qualitative analysis of empirical data is employed.

I claim that the construction of walls in Europe, meant as a means of better security and protection of people within a European state, finally promotes a 'walled in' identity. The isolation of one state makes local residents' mentality also kind of 'walled in': limited or no communication with peoples of other cultures make us view them not just as 'others' but more like 'aliens', threatening our happy homes.

The empirical data was compiled from various sources meeting the aims of the project which includes several stages:

1. the semantic web for the name 'wall' was created to single out the main conceptual characteristics of the lexical units 'wall', 'walled off' and 'walled in' (synonyms and antonyms were singled out from online thesaurus and dictionaries);
2. the corpus of media texts from world wide web with the named lexical units used was compiled, the texts studied are of the period 2016 – 2017;
3. an association experiment was conducted: asking respondents (Russian, German and English native speakers) to put down their first association with the names studied, with special emphasis on emotions and feelings.

To single out main contexts of usage of the unit 'wall' and its synonyms with conceptual meaning of 'defense', 'protection', 'security' etc. Concordance, as a reliable text analysis tool, was used.

The corpus of media texts analyzed contains more contexts with positive estimation of walls which are considered relevant in contemporary world to get protection from outsiders. The walls are means as physical objects. Such contexts often do not openly state that the 'walled off' approach leads to intolerance, hatred, etc. The 'mental' consequence in such cases is the 'walled in' mentality and, thus, isolated, walled in identity.

There are also contexts where 'wall' is meant as metaphor for better protection but not was physical object which leads to logical conclusion that protection should not imply separation and isolation from other cultures. In such cases the mentality of an individual is 'open to the world' which reveals the identity of tolerance and respect to other cultures.

The analysis of associations shows differing views of respondents on the notion of 'walled off'.

On dominating principle of knowledge representation and meaning construction in discourse

Nikolay Boldyrev

Center for Cognitive Research, Derzhavin Tambov State University, Tambov, Russia

Knowledge representation and meaning construction in discourse is always situated and presents a cooperative event. The relationship between knowledge of the world and language use is indirect and depends on how language speakers define it (van Dijk 2009). Obvious as it may seem, this issue still lacks profound insight into the conceptual aspects of verbal interaction and needs consideration of *conceptual factors which involve negotiation of meanings within contexts of knowledge.* The approach to communication based on the importance of factors that surround a communication event has so far remained limited. It has been suggested that in the process of verbal interaction participants rely on certain assumptions that govern conversations in everyday life. In their most widely known form, these assumptions have been expressed as four maxims by P. Grice (see Grice 1989). They jointly specify the so-called "cooperative principle of conversation" within which instances of miscommunication can be analyzed.

In the talk, we argue that the fundamental principle that underlies collaboration among participants is Interpretation Interaction Principle which involves conceptual accommodation, interpretation and negotiation of meanings within contexts of collective and individual knowledge. This principle is based on the concept that there are many workable ways by which individuals can construct their world (Kelly 1963) that is also substantiated in The Linguistic Interpretation Theory (Boldyrev 2016). While contexts of collective knowledge comprise overall knowledge of the world, contexts of individual knowledge reflect the "modification" of the collective knowledge that is influenced by sociocultural parameters, such as the territory the speaker occupies, the education the speaker possesses, the speaker's age, occupation, gender, etc. The word *university*, for instance, represents the following contexts of collective knowledge: "an institution of higher learning with teaching and research facilities typically including a graduate school and professional schools that award master's degrees and doctorates and undergraduate division that awards bachelor's degrees"; "the buildings and grounds of such an institution"; "the body of students and faculty of such an institution" (http://www.thefreedictionary.com/university). In the process of language use, however, this word activates the individual knowledge of a particular speaker: for the driver it activates 'a point in space' (*Can I park at the University?*); for the architect – 'a piece of art' (*The University is in Gothic style*); for common citizens – 'a building, place of work' (*The University Library*), for the child – 'extra activities' (*I take drawing classes at the University*), etc.

Within this view, the problems of meaning construction and language use are discussed in terms of: a) how conceptual systems of different participants are structured and correspond culturally; b) what constitutes the content of conceptual systems; c) whether adequate evaluations of conceptual systems of interactants match; d) the degree of linguistic competence and linguistic performance; e) the mechanisms and cognitive construals that underlie language use (see also: Boldyrev 2017). Hence the research question is to identify variables of conceptual factors that affect language use and meaning construction in discourse. Empirical evidence is mainly drawn from everyday conversations, SMS-messages, blogs, newspaper reports, contemporary books and films, etc.

References:

Boldyrev, Nikolay. 2016. Cognitive schemas of linguistic interpretation. In: Voprosy Kognitivnoy Lingvistiki, 4, 10-20. (In Russ.).

Boldyrev, Nikolay. 2017. Problems of verbal communication in cognitive perspective. In: Voprosy Kognitivnoy Lingvistiki, 2, 5-14. (In Russ.).

Grice, Paul. 1989. Studies in the Ways of Words. Cambridge, Massachusetts: Harvard University Press.

Kelly, George. 1963. A theory of personality: The Psychology of Personal Constructs. New York: The Norton Library.

van Dijk, Teun. 2009. Society and Discourse: how Social Contexts influence Text and Talk. Cambridge: Cambridge University Press.

Acknowledgements: The research is financially supported by the Russian Science Foundation grant at Derzhavin Tambov State University.

Derzhavin Tambov State University, Internatsionalnaya str., 33, Tambov, 392000, Russia boldyrev@tsutmb.ru

Parsing complex sentences in ASD-STE100 within ARTEMIS

Marta González Orta
María Auxiliadora Martín Díaz
Instituto de Lingüística Andrés Bello
Universidad de La Laguna

One of the main objectives of Natural Language Processing (NLP) is the simulation of natural language understanding. Within the different applications designed for this purpose, the ARTEMIS prototype follows the paradigm of unification grammars (Sag, I, Wasow, T. & Bender, E. 2003) and is, at the same time, linguistically grounded in Role and Reference Grammar (RRG - Van Valin & LaPolla 1997 and Van Valin 2005). The syntax-to-semantics linking algorithm proposed in this functional grammar lies at the basis of a parsing process that 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 (LSC) in RRG.

Within ARTEMIS, the semantic component is complemented by FunGramKB, a lexico-conceptual modular knowledge base that consists of an abstract conceptual module comprising an ontology, a cognicon and an onomasticon, and a linguistic module made up of a language-specific lexicon and grammaticon (Periñán-Pascual and Mairal Usón 2011).

A fundamental component in our parser is the Grammar Development Environment (GDE) where production rules (syntactic, lexical and constructional) are stored. Syntactic rules that account for phrasal constituents and simple sentences have already been described in Cortés-Rodríguez and Mairal 2016, Cortés-Rodríguez 2016; Martín Díaz 2017, Diaz Galán and Fumero Pérez 2015, Fumero Pérez and Díaz Galán 2017. This paper will focus, therefore, on the study of complex sentences.

In an attempt to validate these syntactic rules and to avoid some of the common problems that may arise in parsing applications, our research will concentrate on the analysis of RRG's juncture-nexus combinations (Van Valin & LaPolla 1997) and Van Valin (2005) as found in a Controlled Natural Language (CNL), ASD-STE100 (January 2017).

Departamento de Filología Inglesa y Alemana. Campus de Guajara s/n. Universidad de la Laguna 38071-Tenerife.
maorta@ull.es; mmartind@ull.edu.es

The syntactic parsing of ASD-STE100 adverbials in ARTEMIS

Francisco José Cortés Rodríguez
Instituto Universitario de Lingüística 'Andrés Bello'

Universidad de La Laguna

Carolina Rodríguez Juárez
Universidad de Las Palmas de Gran Canaria

This presentation seeks, firstly, to offer an update of the syntactic representation of adverbials in the LCM and FungramKB and, secondly, to implement the conditions necessary for an effective parsing of such constituents within ARTEMIS.

The LCM syntactic representations of sentences are primarily based on the Layered Structure of the Clause (LSC) as proposed in Role and Reference Grammar (Van Valin & LaPolla 1997, Van Valin 2005, Pavey 2010) with some variations motivated by the integration of constructional structures. With regard to the status of adverbials in the LSC and, concomitantly, in the LCM it is striking that they have been sidelined. Proof of this comes from the fact that even though there is a programmatic proposal in Van Valin 2005 for a distribution of adjuncts along the different layers in the LSC, no further contribution has been offered to fully expand such new proposal. Thus, our first aim is to embody the layering proposal for adverbials.

Once this is attained, we also seek to adapt our proposal to the conditions imposed by the Grammar Development Environment (GDE) in ARTEMIS ("Automatically Representing Text Meaning via an Interlingua-Based System"; Periñán-Pascual 2013, Periñán-Pascual & Arcas-Túnez 2014), an NLU prototype developed with the aim of binding natural language fragments with their corresponding grammatical and semantic structures. Since the computational workability of ARTEMIS is still to be tested, we have chosen to apply it on to a Controlled Natural Language, namely ASD-STE100, with the assumption that it will help to validate the performance of our parser. Hence, the scope of our analysis will be confined to the catalogue of adverbials that can be found in a corpus written in this controlled language.

Departamento de Filología Inglesa y Alemana. Campus de Guajara s/n. Universidad de la Laguna 38071-Tenerife. fcortes@ull.edu.es; carolina.rodriguez@ulpgc.es

A Sociolinguistic Corpus Based Investigation of Irish Sign Language Grammatical Classes

Robert Smith
Institute of Technology Blanchardstown Dublin

This paper provides an overview, and some preliminary findings from a sociolinguistic analysis of the Signs of Ireland (SOI) corpus (Leeson et al., 2006).

The work, accomplished using the ELAN media annotation tool (Brugman, 2004) and its various search features, investigates the demographic diversity available in the SOI corpus. The recent addition of grammatical-class transcriptions has for the first time, provided a window into how different social variables affect Irish Sign Language (ISL) performance. It is possible to measure, for example, how often a mature signer uses directional verbs against the frequency of use by signers from younger generations. Given the diverse historical and social context in which ISL has developed (Leeson and Saeed, 2012, pp. 28-57), much can be learned about the language from such sociolinguistic data.

This work presents a methodological approach and some preliminary findings of a wider research project, which explores the form and function of non-manual features in Irish Sign Language (ISL).

The original contribution of this work comes in the form of a corpus-based, sociolinguistic analysis of Irish Sign Language. Carried out, for the first time, with grammatical class transcription data. Furthermore, this work provides a practical contribution in further developing the SOI corpus, such that, the resource is more comprehensively equipped for future research projects.

References

Brugman, H., Russel, A. (2004). Annotating Multimedia/ Multi-modal resources with ELAN. In: Proceedings of LREC 2004, Fourth International Conference on Language Resources and Evaluation.

Leeson, L., Saeed, J., Macduff, A. & Byrne-Dunne, D., 2006. Moving Heads and Moving Hands: Developing a Digital Corpus of Irish Sign Language. Carlow, Ireland, Information Technology and Telecommunications, pp. 25-26.

Leeson, L. and Saeed, J.I., 2012. Irish Sign Language: A cognitive linguistic account. Edinburgh University Press

ABSTRACTS – 6TH JULY, FRIDAY

How can one evaluate a conversational software agent framework?

Dr Kulvinder Panesar

School of Art, Design and Computer Science, York St John University, York, UK

This paper presents a critical evaluation framework for a linguistically orientated conversational software agent (CSA) (Panesar, 2017). The CSA prototype investigates the integration, intersection and interface of the language, knowledge, and speech act constructions (SAC) based on a grammatical object (Nolan, 2014), and the sub-model of belief, desires and intention (BDI) (Rao and Georgeff, 1995) and dialogue management (DM) for natural language processing (NLP). A long-standing issue within NLP CSA systems is refining the accuracy of interpretation to provide realistic dialogue to support the human-to-computer communication.

This prototype constitutes three phase models: (1) a linguistic model based on a functional linguistic theory – Role and Reference Grammar (RRG) (Van Valin Jr, 2005); (2) Agent Cognitive Model with two inner models: (a) knowledge representation model employing conceptual graphs serialised to Resource Description Framework (RDF); (b) a planning model underpinned by BDI concepts (Wooldridge, 2013) and intentionality (Searle, 1983) and rational interaction (Cohen and Levesque, 1990); and (3) a dialogue model employing common ground (Stalnaker, 2002).

The evaluation approach for this Java-based prototype and its phase models is a multi-approach driven by grammatical testing (English language utterances), software engineering and agent practice. A set of evaluation criteria are grouped per phase model, and the testing framework aims to test the interface, intersection and integration of all phase models and their inner models. This multi-approach encompasses checking performance both at internal processing, stages per model and post-implementation assessments of the goals of RRG, and RRG based specifics tests.

The empirical evaluations demonstrate that the CSA is a proof-of-concept, demonstrating RRG's fitness for purpose for describing, and explaining phenomena, language processing and knowledge, and computational adequacy. Contrastingly, evaluations identify the complexity of lower level computational mappings of NL – agent to ontology with semantic gaps, and further addressed by a lexical bridging consideration (Panesar, 2017).

Keywords: conversational software agents, natural language processing, speech act construction, knowledge representation, belief-desire and intentions, functional linguistics

References

COHEN, P. R. & LEVESQUE, H. J. 1990. Intention is choice with commitment. *Artificial Intelligence,* 42**,** 213-261.

NOLAN, B. 2014. Constructions as grammatical objects : A case study of prepositional ditransitive construction in Modern Irish. *In:* NOLAN, B. & DIEDRICHSEN, E. (eds.) *Linking Constructions into Functional Linguistics: The role of constructions in grammar.* Amsterdam/Philadelphia: John Benjamins Publishing Company.

PANESAR, K. 2017. *A linguistically centred text-based conversational software agent.* Unpublished PhD thesis, Leeds Beckett University.

RAO, A. S. & GEORGEFF, M. P. BDI Agents - From Theory to Practice. ICMAS, 1995. 312-319.

SEARLE, J. R. 1983. *Intentionality: An essay in the philosophy of mind*, CUP

STALNAKER, R. 2002. Common ground. *Linguistics and philosophy,* 25**,** 701-721.

VAN VALIN JR, R. D. 2005. *Exploring the syntax-semantics interface*, CUP

WOOLDRIDGE, M. 2013. Intelligent Agents. *In:* WEISS, G. (ed.) *Multiagent Systems.* 2nd edn. ed. USA: Massachusetts Institute of Technology.

School of Art, Design and Computer Science, York St John University, k.panesar@yorksj.ac.uk

Detection of cyber bullying using text mining

David Colton
IBM, Dublin, Ireland
Markus Hoffmann
Department of Informatics, Institute of Technology Blanchardstown, Dublin, Ireland

The internet technology boom has led to a proliferation of tablets, laptops and smart phones with high-speed internet access. This access, coupled with the advent of instant messaging, chat rooms and social media websites, has led to an internet generation who think nothing of posting selfies, mood updates, their relationship status or anything about their life on-line. The traditional bully was the kid in school, or office worker, who got pleasure from watching their victims suffer as they verbally abused them or perhaps made fun of them or maybe even threatened them with physical violence. At least the victim knew who the bully was and, although not a solution, could plan their day to avoid crossing paths with the bully and having to suffer further torment. However, the bully has now also moved on-line. This cyberbully now has twenty-four hour access to a potentially unlimited number of victims. Through their mean and harassing posts and comments the consequences of their cyberbullying activity is too often read about in the papers following another tragic teen suicide. To prevent this new form of bullying, it is important that technology is used to detect these cyberbullying posts.

This paper shows that Python, together with the application of text mining techniques, can be successfully used in the automatic detection of cyberbullying text. The contributions of this paper are many. A new classified cyberbullying dataset, including detailed descriptions of the criteria used in its classification, is generated. An in-depth analysis of several classifiers is undertaken before a novel way of determining the best overall classifier using the recall values of both the positive and negative class is suggested. Finally, an evaluation of the best models is performed by simulating their evolution as new, previously unseen, samples are classified and then included as training data for subsequent iterations.

A qualitative analysis of the wikipedia n-substate algorithm's enhancement terms

Kyle Goslin
Markus Hofmann
Department of Informatics, Institute of Technology Blanchardstown, Dublin 15, Ireland.

Automatic Search Query Enhancement (ASQE) is the process of taking a user submitted search query and identifying terms that can be added or removed to enhance the relevance of documents retrieved from a search engine [1]. ASQE differs from other enhancement approaches as no human interaction is required. ASQE algorithms typically rely on a source of a priori to aid the process of identifying relevant enhancement terms. As Wikipedia contains over 5.5 million articles that are routinely updated, it has been shown to be effective as a source of candidate enhancement terms for ASQE [1].

This paper describes the results of a qualitative analysis of enhancement terms generated by the Wikipedia N-Substate Algorithm (WNSSA) [2] for ASQE. The WNSSA utilises Wikipedia as the sole source of a priori during the query enhancement process. As each Wikipedia article typically represents a single topic, during the enhancement process of the WNSSA, a mapping is performed between the user's original search query and Wikipedia articles relevant to the query. If this mapping is performed correctly, a collection of potentially relevant terms and acronyms are accessible for ASQE. However, some candidate terms may be conceptually distant from the original search query intent, unintentionally impacting the overall performance of the query. This places an emphasis on how relevant each individual term is and the potential available to disrupt the relevance of returned search results.

During a previous analysis of the WNSSA [2], a benchmark was performed using the TREC-9 Web Topics [3] on the ClueWeb12 data set [4]. For each tested search query after enhancement the Average Precision (AP) @ 10 was calculated providing an insight into the relevance of results returned. Although the AP provides an indicator of the overall performance, it does not gauge the quality of enhancement terms nor their individual relevance to the original search query.

This paper reviews the results of the qualitative analysis process performed for each individual enhancement term generated for each of the 50 test search queries. The contributions of this paper include 1) a qualitative analysis of generated WNSSA search query enhancement terms and 2) an analysis of the concepts represented in the TREC-9 Web Topics, detailing interpretation issue during query-to-article mapping performed by the WNSSA.

Keywords: Search Query Enhancement, Text Analysis, Wikipedia.

References

[1] Goslin, K., Hofmann, M., A Wikipedia powered state-based approach to automatic search query enhancement,
Information Processing & Management, 2017.
[2] Goslin, K., Hofmann, M., 2017. A Comparison of Automatic Search Query Enhancement Algorithms That Utilise Wikipedia as a Source of A Priori Knowledge. In *Proceedings of the 9th Annual Meeting of the Forum for Information Retrieval Evaluation* (FIRE'17).
[3] TREC 9 Web Track - https://trec.nist.gov/data/t9.web.html [Accessed March 2018].
[4] ClueWeb12 Dataset - https://lemurproject.org/clueweb12/ [Accessed March 2018].

Department of Informatics, Institute of Technology Blanchardstown, Dublin 15. Kyle.goslin@itb.ie, Markus.Hofmann@itb.ie

Feeding the Lexical Rules in ARTEMIS for the parsing of ASD-STE100

María del Carmen Fumero Pérez
Ana Díaz Galán
Instituto Universitario de Lingüística 'Andrés Bello'
Universidad de La Laguna

There is already a number of contributions (Cortés-Rodríguez and Mairal 2016, Cortés-Rodríguez 2016; Martín Díaz 2017, Diaz Galán and Fumero Pérez 2015, Fumero Pérez and Díaz Galán 2017) which deal with the development of the rules which will embody the G(rammar) D(evelopment) E(nvironment), one of the components of ARTEMIS ("Automatically Representing Text Meaning via an Interlingua-Based System"; Periñán-Pascual 2013, Periñán-Pascual & Arcas-Túnez 2014), a Natural Language Processing prototype aiming at providing the underlying syntactico-semantic representation of linguistic fragments. However, there are still some crucial aspects to be addressed concerning the role played by function words in the development of such rules. Let us recall that the syntactic analysis offered by the GDE involves a significant variation with regard to the original LCM syntactic apparatus, which fed from the Role and Reference Grammar Layered Structure of the Clause (Van Valin & LaPolla 1997, Van Valin 2005, Pavey 2010). The most important changes affect the operator projection, which is overridden by the so-called Attribute-Value Matrixes (AVMs), a theoretical construct borrowed from unification grammars (as f.i. Sag, I, Wasow, T. & Bender, E. 2003, among others).

Since function words are heavily responsible for the encoding of grammatical information of the kind represented originally by operators, their role is vital in the parsing process carried out by the GDE. The aim of this presentation is to offer a detailed account of the Lexical Rules necessary for the integration of function words in ARTEMIS. Lexical rules are the means to encode the relevant morphosyntactic information attached to each functional item which later will be integrated into the higher syntactic structures where they participate.

In its present state of development ARTEMIS is being implemented for a controlled natural language, ASD-STE100, under the assumption that its simplified nature will help pave the way to the eventual parsing of a natural language. Our research will therefore concentrate on feeding the lexical rules necessary for the analysis of this simplified language.

Departamento de Filología Inglesa y Alemana. Campus de Guajara s/n. Universidad de la Laguna 38071-Tenerife.
mfumero@ull.edu.es; adiazgal@ull.edu.es

Functional-semantic status of lexical-grammar parenthesis-modal discourse-text «transitions» in modern English and French languages

Sabina Nedbailik

Institute of foreign languages, Petrozavodsk State University, Petrozavodsk, Russia

As it is known, discourse-text lexical-grammar parenthesis-modal «transitions» present a non ordinary, complex, multi-aspect, hard for perception in the frame of oral and written speech models language phenomenon, having hybrid functional-semantic characteristics and not homogeneous genesis in modern English and French languages. Existing at the junction of auxiliary and full-semantic lexical systems, the elements of this group show a tendency for interaction, i.e. integration/differentiation with/from adverbs and substantives, as well as so called pure copulas – primary bearers of conjunction category semantics. In this connection a special attention is paid to pointing out and describing the basic aspects, factors, particularities and prospects of their integration/differentiation with various auxiliary and full semantic lexical units.

The results of practical material complex, detailed and multi-aspect structural, comparative analysis carried out in the frame of our research give the possibility to state a large functional-semantic scope and lexical-grammatical features hybrid, mixed character of most parenthesis-modal adverbial-substantive, verbal and other types discourse-text connectors both in English and French languages, what enables them to play the role of lexical -semantic intensifiers as well as discursive modal markers, attenuators in the frame of whole statements, syntactic complexes, structures and (super)phrasal unities, having frequent occasional or fixed use both on sentential and inter-sentential levels. Of course, high migration abilities of most lexical-grammar discourse-text parenthesis-modal «transitions» explain their considerable transposition and transformation potential.

Considering their characteristics polyphony, discourse-text connectors of various genesis and type are differentiated by stylistic universality/ specialization, occasional/usual character of functioning, what is explained by their basic specific semantic features. In this connection, the most important for purposes of utterances syntactic structuring, segmentation, composing and highly interesting in the aspect of functional-stylistic variations, in our opinion, are the logical subtype «transitions» of adverbial, substantive, verbal, adjective origin, expressing correspondingly meanings of:

1) precedence (En.) to begin with, to start, first, primarily, etc.; (Fr.) premièrement, tout d`abord, pour commencer, etc.)
2) reason ((En.) because of this, as, since, etc.; (Fr.) en raison de, grâce à, à force de, etc.)
3) consecution/consequence ((En.) as a result, finally, to sum up, etc.; (Fr.) finalement, par conséquent, enfin, c`est pourquoi, etc.)
4) addition ((En.) to add, besides, moreover, etc.; (Fr.) en/de plus, en outre, non seulement..., etc.)
5) contrasting ((En.) still, by contrast, however, etc.; (Fr.), malgré, par contre, cependant, etc.)
6) emphasizing ((En.) really, surely, in fact, etc.; (Fr.) en effet, justement, juste, etc.),
7) illustration ((En.) for example, to be exact, etc.; (Fr.) d`après, selon, prenons le cas de…, ainsi, etc.)
8) comparing ((En.) similarly, likewise, at the same time, etc.; (Fr.) autrement dit, (non) moins (que), plus, etc.)
9) generalizing ((En.) to sum up, on the whole, in general, etc.; (Fr.) en somme, généralement, bref, etc.)
10) concretizing (En.) in particular, particularly, etc.; (Fr.) par example, c`est à dire, etc.)

Of course, lexical-grammar parenthesis-modal «transitions» using frequency degree in both modern English and French languages greatly depends on discourse-text utterances style and gender characteristics.

Institute of foreign languages. Petrozavodsk State University, Petrozavodsk, 185680. german@petrsu.ru

The role of persuasion processes in shaping various methods of message framing

Anna Kuzio

University of Zielona Gora, Department of Humanities, Zielona Gora, Poland

In spite of the large quantity of research devoted to the aspect of persuasive messages that aim at presenting behavior and attitude modification, resources are not exploited on messages that are eventually feeble. In some instances, theoretically persuasive messages may be unproductive since the message may not highlight content that the precise message recipient seems to be most receptive to. Moreover, there seems to be some possibly that persuasive messages may be unsuccessful as some communicators aim to target an attitude in the future, some types of messages only shape attitudes for a short duration of time. Eventually, comprehending how some types of persuasive messages function can help communicators exploit them efficiently.

Message framing appears to be one method of persuasion that can be an effective device in persuading individuals to adapt viewpoints and to assess issues more favorably (Lee & Higgins, 2004). Message framing tends to be seen as a way of manipulating characteristics of a message to be well-matched with the way individuals naturally view goals (Higgins, 1997). Some individuals are inclined to look at goals as expectations and aspirations and some individuals have a tendency to view goals as duties and obligations. This goal orientation is called a regulatory focus and is assumed to affect how individuals respond to constructed messages (Higgins, 2000). Dependent on one's regulatory focus, message features such a highlight on positive outcomes that happen as a result of adapting a suggested behavior (gain frame), the stress on negative results that happen as a result of not adapting a recommended behavior (loss frame), the stress on positive cues in a message (promotion focus), and the stress on negative cues in a message (prevention focus) can decided how persuasive a message is.

It is the aim of the present study to examine how different methods of message matching function. Precisely, the present study combines traditional self-report measurement questionnaires with physiological eye tracking data to check if all the methods of matching will trigger the effects of regulatory fit such as feelings of fluency, feeling right, and positive attitudes- matches that include dispositional regulatory focus will cause more visual attention to and more discerning about the message content, supporting the general hypothesis that message match that comprises dispositional regulatory focus contain more attention to and thinking about the message content.

References:

Higgins, E.Tory. 1997. Beyond Pleasure and Pain. American Psychologist, 52(12), pp. 1280-3000.

Higgins, E.Tory. 2000. Making a good decision: Value from fit. American Psychologist, 55, pp. 1217-30

Lee, Angela Y., Aaker, Jennifer L. 2004. Bringing the frame into focus: The influence of 40 regulatory fit on processing fluency and persuasion. Journal of Personality and Social Psychology, 86, pp. 205-218.

Department of Humanities, University of Zielona Gora, Zielona Gora, Poland. annakuzio@gmail.com

Metaphor-facilitated co-creation strategy in election campaigns

Inna Skrynnikova
Institute of Philology and Intercultural Communication,Volgograd State University, Volgograd, Russia

One of the key questions of the interdisciplinary studies in recent years is how digitization impacts the world around us and affects the ways people see critical societal issues. The use of co-creation in online social (health, marketing, traffic safety) and political (election, foreign affairs, state of the economy) campaigns is one of these challenges. Unlike classical campaigns which are typically top-down with organizations being responsible for the entire campaign, co-creation campaigns are a special type of collaboration between organization and target audience, facilitated through Web 2.0 technologies. In this sense such campaigns can be the subject of research in many disciplines bringing together communication science (i.e., social science) and linguistics (i.e., humanities).

Carefully worded and well-organized current election campaigns employ numerous persuasive strategies to make voters take a particular political stance. One of the most recent and powerful ones is the web-based co-creation strategy unfolding in political blogs and social networks and enabling the audience members (prospective voters) feel they are active contributors to the campaign (Zwass, 2010). There is mounting evidence from other fields (social campaigns, advertising, creative writing) which confirm the efficiency of this strategy in achieving a goal to be sought (Heerik Burgers et al. 2017). The essence of the co-creation strategy is to ask audience members through social media to complete a political slogan, previously exposing it to a particular type of metaphor. Even if most audience members follow the campaign's directive, some might deviate and write a message contradicting a negative view promoted by co-creation campaign organizers (e.g., LIVING IN ISOLATION is PROMOTING DOMESTIC INDUSTRIES). Such ambivalent responses can be also found for many co-creation campaigns (e.g., Gebauer et al., 2013; Heidenreich et al., 2015), which leads us to the question when and how co-creation campaigns succeed or fail.

The current paper takes on the challenge to answer the question by providing new ground in two ways critical for its understanding. The first one is to present an interdisciplinary perspective by exploring the role of both cues of the online political environment (*social and political sciences*) and linguistic cues (*linguistics*) in revealing the ways in which audience members can co-create election campaign slogans. The second one is to exemplify that, many co-creation initiatives can fail or succeed depending on the strategies employed.

Although these seem to provide deeper insights into the nature of the strategy, it is still difficult to establish which factors are responsible for audience contributions to creating slogans. Thus, we propose to supplement such studies with a production experiment in which respondents co-create slogans in a controlled research environment. Specifically, we randomly exposed participants to one of several co-creation campaigns that differ in cues in the online environment and in linguistic metaphor-based cues to trigger co-creation. Such a procedure in our view can contribute to our understanding of the ways these cues work together in generating audience responses. It can also further provide ambiguous evidence as to which elements constitute successful web-based co-creation strategies in election campaigns.

References

Durugbo, C., & Pawar, K. (2014). A unified model of the co-creation process. *Expert Systems with Applications*, *41*(9), 4373-4387.

Heidenreich, S., Wittkowski, K., Handrich, M., & Falk, T. (2015). The dark side of customer co-creation: exploring the consequences of failed co-created services. *Journal of the Academy of Marketing Science*, *43*(3), 279-296.

McAdams, D. P., Albaugh, M., Farber, E., Daniels, J., Logan, R. L., & Olson, B. (2008). Family metaphors and moral intuitions: How conservatives and liberals narrate their lives. *Journal of Personality and Social Psychology*, *95*(4), 978-990.

Steen, G. J. (2011). The contemporary theory of metaphor: Now new and improved!. *Review of Cognitive Linguistics*, *9*(1), 26-64.

van den Heerik, R.A.M., van Hooijdonk, C.M.J., Burgers, C., & Steen, G.J. (2017). "Smoking is sóóó… sandals and white socks". Co-creation of a Dutch anti-smoking campaign to change social norms. *Health Communication.*

Veale, T. (2012). *Exploding the Creativity Myth: The Computational Foundations of Linguistic Creativity*. London: Bloomsbury.

Institute of Philology and Intercultural Communication, Volgograd State University, Volgograd, 400062.
i.skrynnikoval@volsu.ru

The forms, functions and pragmatics of Irish polar question–answer interactions

Brian Nolan
Institute of Technology Blanchardstown

In this paper we examine the challenges of unpacking meaning and characterising knowledge in the speech act of requesting information in one of its manifestations, the polar yes-no question, for Irish. Irish does not have any exact words which directly correspond to English 'yes' or 'no' and so employs different strategies where a yes-no answer is required. We characterise the expressive forms, functions, logical underpinnings and pragmatics of polar yes-no interrogatives as question–answer pairs, and the felicity conditions necessary for their successful realisation. In a question-answer interaction, information is assumed to be freely exchanged, under a Gricean presumption of cooperation.

A polar yes-no question in Irish can be considered as advancing a hypothesis for confirmation and consequently, there are several strategies available for answering a polar yes-no question. In Irish, the answers to yes-no questions echo the verb of the question for both affirmative and negative answers, along with a negation marker for negative answers. These types of answers are referred to as verb-echo answers. Typically, In Irish, the verb form is used *without* explicit nominal arguments expressed within grammatical relations, though there are exceptions. Additionally, in negative polarity answers, the negative particle is also used. When a synthetic verb form is used, a pronominal appears in the grammatical relation of nominative subject within the answer. In the case of analytic verb forms, the subject is always missing. A subject is used only when the speaker chooses an emphatic affirmation or denial. The verb within the answer is inflected for tense as well as subject agreement. As tense is a clausal operator, it locates the time of the event denoted by a clause in relation to the time of utterance. The presence of tense in the answer implies the presence of a clause. When it occurs, pronominal subject marking implies the presence of a subject, hence also the presence of a clause. Under certain circumstances, as the answer to a copula-question with an indefinite predicate, the copula-derived phrases *sea* (*COP+3SG* = 'be-it') and *ní hea* (= NEG.COP *3SG* 'NEG be it'), function as logically equivalent to 'yes' and 'no'.

This paper argues towards several claims regarding polar yes-no questions of Irish. One claim is that the answers to polar yes-no questions of Irish contain instances of ellipsis and, as such, represent full clausal expressions with a complete semantics where the elided elements are from the question part of the question-answer pair. The propositional content is inferred from the context, specifically from the question with which the answer is paired. Another claim is that one of the functions of interrogatives is the maintenance of common ground via the update and exchange of information between the interlocutors. It also serves to reinforce social affiliation in a group through having access to shared knowledge and understanding. The fact that languages have clausal types for requesting information, and asking (polar yes-no) questions, shows clearly how important this activity is to human communication, and the construction and maintenance of common ground, and meaning.

A group theory for conceptual meanings (Digital Linguistics)

Kumon Tokumaru
Oita, Japan E-mail: kalahari145ma@y-mobile.ne.jp

Digital Linguistics is an interdisciplinary study that identifies human language as a digital evolution of mammal analog vocal sign communications, founded on the vertebrate spinal sign reflex mechanism. Analog signs are unique with their physical sound waveforms but limited in number, whilst human digital word signs are infinite by permutation of their logical property, phonemes.

Figure-1 demonstrates an overview of linguistic intelligence. Linguistic information takes the shape of noise-resistant mono-dimensional structures of discrete and known signals, i.e. sequence of syllables, in the physical layer. On closer examination, there are conceptual and grammatical syllables, and it is clear that each grammatical syllable is modulating its adjacent concept. It seems that speech sound consists of a minimum semantic unit of a concept and grammar, and the brain processes linguistic information unit by unit, which indicates that the complexity lies in concepts which develop inside the individual brains.

Figure-1 Linguistic Intelligence Analysis based on OSI Reference Model with Noise Factors

	Teacher's brain		Disciple's brain
Applications	Human Collective Intelligence	DYNAMIC and EVOLVING	Further Development of Intelligence
Presentation	Index, Chapters, References	PHENOMENA	Reindex, Recombine:
Session	Method Attitude,	in Extremely Low Noise	Rebuild: New word memory generation
	Key Concepts,	Logical Layer Environment	Dualistic thought operations
	Chronology, (Betrayal)	i.e. Ventricle System, CSF,	Interdisciplinary integration
Transport	Supralaryngeal Vocal Tract,	Brainstem Reticular Formation,	Receive: Ag/Ab, Data Verification:
	Data Protection Measures (coding)	in Monasteries, Temples	Aucenticity, Forward Error Correction
	From Brain to Speech Voice	DISCRETE, KNOWN, LINEAR	From Speech Voice to Brain
Network	Face-to-Face Lecture,	NOISE RESISTANT FORM	Translating foreign/ancient languages,
	Secretary/Clerks, Publisher	(= Mono-Dimensional)	Literacy (ability to vocalize written text)
Data Link	Printing Transcription Copy	in Noisy Physical Layer	Search engine, OPAC, citation by others
Physical	Speech, Manuscripts, Books, Archives, Electric/Electronic Recordings, Database, Archive		

NOISE FACTORS	
Logical	Falsifications, Apocryphal, Censorship
Physical	War, Fire, Water, Chemical

At MKR6, the author submitted a new hypothesis on the brain mechanism for linguistic processing, where B-lymphocytes inside the CSF are identified as conceptual devices equipped with the logic of dichotomy and dualism. [Jerne1974] [Jerne1984] [Tokumaru2017a] [Tokumaru2017b] Dichotomy executes pattern recognition of learned words. (A or not-A) Dualism allows us to formulate versatile logical circuits such as “if A then B” (Reminiscence of memory B), “A+B=C” (Judgment, Evaluation: B= conditions or another concept, C = logical memories such as ◎○×△＝≠<≒) “A+B=C” (Integration, Complication: C = conceptual memories). Thus, meanings of any complex concept can be constructed through the operation of dichotomies and dualisms.

Sensory, logical and conceptual memories connected to a concept consist of a group, which is its meanings. J. Piaget [1947] made a pioneering study on the functional meaning and structure of “groupings”, which are similar to “groups” in mathematics and are operable with five simple formulae of Combinativity, Reversibility, Associativity, General operation of identity and Tortology. These formulae are necessary when we want to identify, examine and verify conceptual meanings.

(I) Combinativity: $x + x^1 = y$; $y + y^1 = z$; etc.

(II) Reversibility: $y - x = x^1$ or $y - x^1 = x$.

(III) Associativity: $(x + x^1) + y^1 = x + (x^1 + y^1) = (z)$.

(IV) General operation of identity:

$x - x = 0$; $y - y = 0$; etc.

(V) Tautology or special identities:

$x + x = x$; $y + y = y$; etc.

As meanings consist of memories, they are based on individual experiences, learning and thought operations. “The remarkable fact in the continuous assimilation of reality to intelligence is, in fact, the equilibrium of the assimilatory frameworks constituted by grouping. ….. Throughout its formation, thought is in disequilibrium or in a state of unstable equilibrium; every new acquisition modifies previous ideas or risks involving a contradiction. From the operational level, on the other hand, the gradually constructed frameworks, classificatory and serial and spatial, temporal, etc., come to incorporate new elements smoothly; the particular section to be found, to be completed, or to be made up from various sources, does not threaten the coherence of the whole but harmonises with it.” [Piaget1947]

This equilibrium, which is not desirable for the development of human intelligence, seems to be a trace of the vertebrate spinal sign reflex in charge of fundamental activities such as food, security, reproduction, etc. In order to overcome restrictions attributable to this reflex, we should keep in mind that sciences are the *raison-d’être* of linguistic humans.

References

Jerne, N.K. (1974) Toward a Network Theory of Immune System, Ann Immunol (Paris). 125C(1-2) :373-89

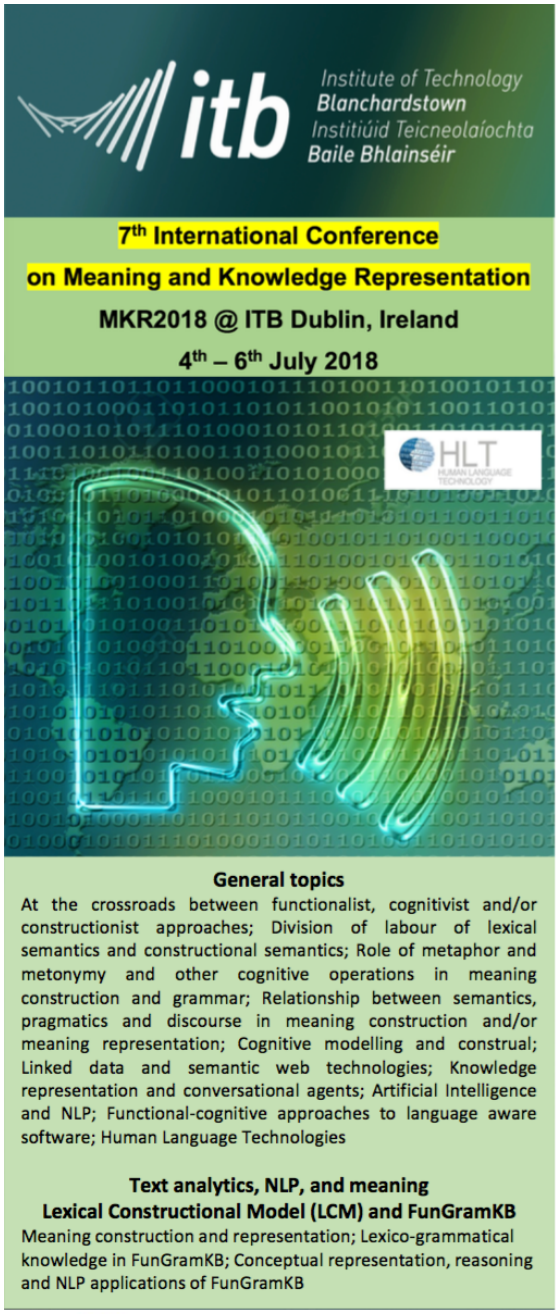
Jerne, N.K. (1984) The Generative Grammar of the Immune System

Piaget,J. (1947) La psychologie de l'intelligence, Paris, Armand Colin.

Tokunmaru K. Inside Ventricle System Immune Cell Networks for the Mechanism of Meaning: B-lymphocytes floating in cerebrospinal fluid (CSF) network with glial cells in neocortex (A Hypothesis), 6MKR Publication (2017a)

Tokumaru K., Psychology, logics and molecular /cellular biology for the mechanism of concept and grammar: inside ventricle system immune cell network hypotheses, Cogn. Studies of Language Vol.XXX Sept 20-22, 2017, Rus. Aca. Sci. pp860-864 (2017b)

=== === End of MKR2018 abstracts === ===



Institute of Technology
Blanchardstown
Institiúid Teicneolaíochta
Baile Bhlainséir