## **BOOK OF ABSTRACTS**



# Meaning and Knowledge Representation

11th International Conference

## University of Almería 3-5 July 2024









AGENCIA ESTATAL DE INVESTIGACIÓN

## Book of abstracts. Meaning and Knowledge Representation 11th International Conference

© del texto: Los autores

Libros Electrónicos n.º 178

© de la edición: Editorial Universidad de Almería, 2024 editorial@ual.es www.ual.es/editorial Telf/Fax: 950 015459 ¤ ISBN: 978-84-1351-320-1



Grant PID2020-112827GB-I00 funded by MICIU/AEI/10.13039/501100011033



pulsando el pie de la página

## Organising committee

## **Chairs:**

María Enriqueta Cortés de los Ríos (Universidad de Almería, Spain) María del Mar Sánchez Pérez (Universidad de Almería, Spain)

## **Committee members:**

Madalina Armie (Universidad de Almería, Spain) Germán Asensio Peral (Universidad de Almería, Spain) Carmen Bretones Callejas (Universidad de Almería, Spain) Ángel Miguel Felices Lago (Universidad de Granada, Spain) Nicolás José Fernández Martínez (Universidad de Jaén, Spain) Rocío Jiménez Briones (Universidad Autónoma de Madrid, Spain) Ángeles de las Mercedes Jordán Soriano (Universidad de Almería, Spain) Fredy Núñez Torres (Pontificia Universidad Católica de Chile, Chile) M. Beatriz Pérez Cabello de Alba (UNED, Spain) Carlos Periñán-Pascual (Universitat Politècnica de València, Spain) José Javier Torres Fernández (Universidad de Almería, Spain)

## Scientific committee:

Francisco Arcas Túnez (Universidad Católica San Antonio, Spain) Annalisa Baicchi (Università di Pavia, Italy) Hans C. Boas (University of Texas at Austin, USA) Chris Butler (Swansea University, UK) Francisco Cortés Rodríguez (Universidad de La Laguna, Spain) Elke Diedrichsen (Institute of Technology Blanchardstown, Ireland) Ángel Felices Lago (Universidad de Granada, Spain) Carlos González Vergara (Pontificia Universidad Católica de Chile) Francisco Gonzálvez García (Universidad de Almería, Spain) Markus Hofmann (Institute of Technology Blanchardstown, Ireland) Rocío Jiménez Briones (Universidad Autónoma de Madrid, Spain) Ricardo Mairal Usón (UNED, Spain) Brian Nolan (Institute of Technology Blanchardstown, Ireland) Klaus-Uwe Panther (Universität Hamburg, Germany) María Beatriz Pérez Cabello de Alba (UNED, Spain) Sandra Peña Cervel (Universidad de La Rioja, Spain) Carlos Periñán-Pascual (Universitat Politècnica de València, Spain) Francisco Ruiz de Mendoza Ibáñez (Universidad de La Rioja, Spain) Robert D. Van Valin, Jr. (University at Buffalo. The State University of New York, USA)

## **Table of Contents**

Ρl	ENARY TALKS
	<b>Charles Forceville:</b> "What Steers the Interpretation of a Visual or Multimodal Message? A Relevance Theory Perspective" (Based on Forceville 2020)
	<b>Francisco Gonzálvez-García:</b> Capturing Meaningful Generalizations at Varying Levels of Resolution: The Case of the Family of " <i>ser</i> + <i>muy</i> d <i>e</i> -PP" Constructions in Spanish <b></b>
	<b>Maite Martín Valdivia:</b> Automatic Generation of Counter-Narratives to Combat Hate Speech Using Language Models10
AI	BSTRACTS
	Ángela Alameda Hernández & Pedro Ureña Gómez-Moreno: Analysing User-Generated Content
	on Social Media to Understand the Social Problem of Violence Against Women
	Bartłomiej Alberski: Structure and Applications of Plwordnet: Analysis of the Macro- and Microstructure of the Largest Polish Relational Dictionary
	<b>Roberto Asenjo Alcalde:</b> Metaphors in the Trenches. Polarization in the Posters of the Spanish Civil War
	Fatima Azzahraa El Yamlahi & María Enriqueta Cortés de los Ríos: Multimodal Metonymy in Branding: a Corpus-Based Account of Moroccan Food Products
	Carmen M. Bretones Callejas: Cognitive Condensation and Synesthetic Force in the Works of Seamus Heaney
	Jana Dietzel: Freudian Metaphor: Analogies of Psychoanalytic Terminology to Science, Space, and Body Concepts
	<b>Ángel Felices Lago &amp; Rocío Jiménez-Briones:</b> Applying Corpus Methodology to Detect Social Problems in User-Generated Content Units
	Nicolás José Fernández Martínez: Guidelines for LLM-Assisted Corpus Building and Annotation
	Alicia Galera Masegosa: The Role of Echoing in the Conventionalization of Implicational Meaning
	<b>Hongling Gou:</b> Enhancing Linguistic Annotation for Emotion Analysis: a Hybrid Method for Computer-Assisted Annotation
	Linxi Jiang: Preliminary Design of an Ontology for the Domain of Orthopedic Diseases
	Svetlana Kiseleva & Nella Trofimova: Principles of Metaphtonymy Formation (Based on English Verbs With Semantics "to Divide")
	Sanjaya Lenka: Cognitive Operation in Conversational Meaning
	Sayf Mohamed: Metaphors and the Construction of Climate Change Discourse. A Corpus Based Study of Multimodal Metaphors
	<b>Carla Ovejas Ramírez &amp; M.ª Sandra Peña Cervel:</b> Lexical Blends in Spanish: an Analysis from the Point of View of Cognitive Linguistics
	<b>Kulvinder Panesar &amp; María Beatriz Pérez Cabello de Alba:</b> Knowledge Graphs Interpretation for the Speech and Language Analysis of a Multimodal Al Agent for Pre-Screening Cognitive Function of Older Adults
	M.ª Sandra Peña Cervel & Carla Oveias Ramírez: The Constructional Nature of Merism 35

Carlos Periñán-Pascual: Synthetic Datasets for Community Problem Detection in ALLEGRO	36
Adrián Raso González: Functional Concepts or Functions and Concepts? Boundaries of the Conceptual-Procedural Distinction and LSA	37
Chenchen Song: Two Levels of Lexical Ambiguity and a Unified Categorical Approach	39
Hamest Tamrazyan & Gayane Hovhannisyan: Translating Cultural Knowledge Representation for Heritage Preservation	41
<b>Cristina Zimbroianu:</b> Context Theory Versus Politeness Theory Applied to the Great Fortune and Spoilt City Novels by Olivia Manning to Refute Romanian Critics and Censors	43

## **Plenary talks**

## **"WHAT STEERS THE INTERPRETATION OF A VISUAL OR MULTIMODAL MESSAGE? A RELEVANCE THEORY PERSPECTIVE" (BASED ON FORCEVILLE 2020)**

Charles Forceville

University of Amsterdam (Netherlands)

Multimodality and semiotics scholarship is in need of an inclusive model of communication that takes into account the identities of the communicator, the audience, as well as their relation, and that does not privilege specific media and/or modes over others. The contours of such a model exist in relevance theory/RT (Sperber and Wilson 1995), whose central claim is that *each act of communication comes with the presumption of optimal relevance to the envisaged audience*. Hitherto RT scholars (typically: linguists) have almost exclusively analysed face-to-face exchanges. To fulfil RT's potential to develop into an inclusive theory of communication, it is necessary to explore how it can be adapted and refined to account for (1) messages in other modes than (only) the verbal mode; and (2) mass-communication. In Forceville (2020) I propose how RT works for mass-communicative messages that involve static visuals. In my presentation I will specifically focus on how RT approaches the key issue of which factors have an impact on the interpretation of a picture or a multimodal message, discussing this issue by drawing on examples from different genres (logos & pictograms, advertisements, and cartoons).

c.j.forceville@uva.nl

## CAPTURING MEANINGFUL GENERALIZATIONS AT VARYING LEVELS OF RESOLUTION: THE CASE OF THE FAMILY OF *"SER + MUY DE-PP"* CONSTRUCTIONS IN SPANISH

Francisco Gonzálvez-García

University of Almería (Spain)

On the basis of 1,710 examples of Spanish es muy de ('is very of') from the Corpus del Español NOW, together with acceptability judgements from 10 native speakers of Castilian and South American Spanish, this paper provides a principled constructionist account (cf. Goldberg, 2006; Goldberg and Herbst, 2021) of the main characteristics of expressions like (1)-(4) below, with special focus on the acceptability differences of the nominal slot in the PP.

(1) Él es muy de bares/del siglo XXI/de un equipo como la Juventus

Lit. 'He is very of bars, /of the century XVI/of a team like the Juventus.'

'He is very into bars/very XXI century/a big supporter of a team like the Juventus.'

(2) Él es muy de levantarse muy temprano

Lit. 'He is very of get.up.himself very early.'

'He is very into getting up early.'

(3) *Mi marido* es *muy de que yo haga lo que me propongan* 

Lit. 'My husband is very of that I should.do the which me propose.'

'My husband is very in favour of me doing whatever I'm asked to do.'

(4) (a) *Tu ayuda es muy de agradecer* 

Lit. 'Your help is very of thank.'

'Your help is very appreciated.'

- (b) Él es muy de **fiar**
- Lit. 'He is very of trust.'
- 'He is very trustworthy.'

Examples (1)-(4) are best handled in terms of coercion between an intensifier (*muy* 'very') and non-stative/non-gradable elements, such as (i) bare nouns, definite and indefinite NPs (cf. (1)), (ii) active infinitival clauses (cf. (2)), (iii) finite *que*- ('that')-clauses (cf. (3)), and passive infinitival clauses (cf. (4)). These combinations qualify as stativizing constructions (Michaelis 2011) and, more specifically, as individual-level predicates (i.e. stative predicates denoting an essential, time-stable property) with a characterizing, evaluative interpretation (Fernández-Leborans and Sánchez López, 2015: 112). This explains why *ser* ('be') cannot alternate with estar ('be.at') in constructions of this kind (cf. (5)), and why PPs resisting a gradable, evaluative construal are infelicitous (cf. (6)(a)-(b)).

(5) *Él está muy #de bares/#de levantarse temprano* 

Lit. 'He is.at very of bars/very of get.up early.'

(6) (a) La temperatura es de **30 grados** 

'The temperature is of 30 degrees.'

#### (b) #La temperatura es muy de 30 grados

Lit. 'The temperature is very of 30 degrees.'

The specific constructional interpretations in (1)-(4) arise from contextual adjustments (Carston, 2015), including variables like (i) the (non-)human nature of the main clause subject, (ii) to extent to which the state of affairs is controllable by the main clause subject, and (iii) the actual or potential iterativity of the state of affairs. Thus, the coerced PPs may be contextually modulated as encoding a person's habits (cf. (1)-(2) or inclinations (cf. (3)), a potential modal deontic habituality (cf. (4)(a)), and/or an evaluative property of a person, entity or event (cf. (4)(b)).

In sum, the semantico-pragmatic hallmarks of the sub-constructions in (1)-(4) can be subsumed under a family of *ser muy de*-PP constructions, with this general meaning: 'X (SOMEONE/SOMETHING) (SUBJECT) IS SUBJECTIVELY CONSTRUED AS HAVING Y (A HIGHLIGHTED CLASSIFICATORY PROPERTY (OF AN INDIVIDUAL/CLASS) (ATTRIBUTE)'.

#### REFERENCES

Carston, R. (2015). Contextual adjustment of meaning. In N. Riemer (ed.), *The Routledge Handbook of Semantics*. London & New York: Routledge, 195-210.

Corpus del Español NOW. https://www.corpusdelespanol.org/now/

- Fernández-Leborans, M. J., and C. Sánchez López. (2015). Sentences as predicates: The Spanish construction <ser muy de + infinitive>. In I. Pérez-Jiménez, M. Leonetti, and S. Gumiel-Molina (eds.), New Perspectives on the Study of Ser and Estar. Amsterdam: John Benjamins, 85-118.
- Goldberg, A. E. (2006). *Constructions at Work: The Nature of Generalization in Language*. New York: OUP.
- Goldberg, A. E., and T. Herbst. (2021). The *nice-of-you* construction and its fragments. *Linguistics*, 9(1), 285-318.

Michaelis, L. (2011). Stative by construction. *Linguistics*, 49(6), 1359-1400.

fgonza@ual.es

## AUTOMATIC GENERATION OF COUNTER-NARRATIVES TO COMBAT HATE SPEECH USING LANGUAGE MODELS

Maite Martín Valdivia

University of Jaén (Spain)

Hate speech has become a global concern, particularly affecting vulnerable people and communities across multiple online platforms. In this talk, we will explore an innovative strategy to tackle this problem, in the framework of the European project RealUP: Combating Hate Speech. Using advanced language models, such as GPT-4 or Mistral, we will demonstrate how it is possible to generate effective counter-narratives that can mitigate the impact of hate speech in real time. The presentation will start with an introduction to the problem, showing the different approaches that have been implemented over time and highlighting how the integration of large language models has rapidly evolved to address this specific challenge. Examples of these approaches will be presented, analysing both their strengths and the challenges encountered. In addition, future work and possible lines of continuation in this research will be discussed, with the aim of improving and extending the effectiveness of automatic counter-narratives.

maite@ujaen.es

## Abstracts

## ANALYSING USER-GENERATED CONTENT ON SOCIAL MEDIA TO UNDERSTAND THE SOCIAL PROBLEM OF VIOLENCE AGAINST WOMEN

### Ángela Alameda Hernández and Pedro Ureña Gómez-Moreno

## University of Granada (Spain)

In the age of digital social networks, their users have become effective providers of instant and authentic information through the content they post online. The detection and analysis of this usergenerated content (UGC) can be extremely useful when such messages deal with issues that affect our society. Its analysis can enable the relevant institutions to take the appropriate actions to tackle specific social problems. In this line, a salient problem that affects physical safety in our society today is gender-based violence and, in particular, acts of violence where women are the victims.

In this context, the present research analyses a corpus of text messages about violence against women from the social media platform Twitter/X. This corpus is integrated within the multimodal intelligent system ALLEGRO, which has been developed for the analysis of social media data and aims to detect problems and adverse events in real time. ALLEGRO consists of different modules for text, image, and sound processing. In this paper, the focus will be on the first of these modules, named DIAPASON, since the corpus under study is a compilation of written UGC units that deal with the social problem of violence against women.

This paper shows the results of the textual analysis of this corpus using the TEXMILAB programme. It is designed as a linguistic laboratory and includes tools for corpus building, text preprocessing, and text analysis. The results show the analysis of unigrams, bigrams, collocations, and concordances, with the aim of identifying the key words and structures that social media users employ in relation to the problem of violence against women, in order to shed light on this social problem and the concepts associated with it. Overall, the analysis presents how the social problem of violence against women is discursively produced and reproduced by ordinary citizens as social media users who generate media content and can have a notable and practical social impact contributing to the understanding of this social problem and how it is perceived and represented in our society.

#### REFERENCES

- Mooney, L. A., Clever, M., & Van Willigen, M. (2021). Understanding social problems. Boston: Cengage learning.
- United Nations Population Fund (2022). "Gender-based violence". <u>https://www.unfpa.org/gender-based-violence</u>
- Panagiotou, N., Katakis, I., Gunopulos, D. (2016). "Detecting Events in Online Social Networks: Definitions, Trends and Challenges". In Michaelis, S., Piatkowski, N., & Stolpe, M. (eds) Solving Large Scale Learning Tasks. Challenges and Algorithms. Springer Cham: 42-84.
- Periñan-Pascual, C. (2023). "Modelización de las quejas de los ciudadanos como artefactos digitales culturales: DIAPASON". In Olmo-Cazevieille, F., *Investigación lingüística en entornos digitales*. Valencia: Tirant Lo Blanch.

Shepherd, L. (2008). Gender, violence, and security: Discourse as practice. London: Zed books.

aalameda@ugr.es; pedrou@ugr.es

## STRUCTURE AND APPLICATIONS OF PLWORDNET: ANALYSIS OF THE MACRO- AND MICROSTRUCTURE OF THE LARGEST POLISH RELATIONAL DICTIONARY

Bartłomiej Alberski

#### Wrocław University of Science and Technology (Poland)

The aim of the paper is to present the macro- and microstructure of the largest Polish relational dictionary - plWordNet (http://plwordnet.pwr.wroc.pl/wordnet/), which is a resource utilized in numerous Natural Language Processing (NLP) applications within the Clarin-PL project. This wordnet is constructed based on the Princeton WordNet but is significantly larger and operates with a greater number of relations (see Table 1). The foundation of Słowosieci as a relational dictionary lies in clustering lexical units into synsets (groups of synonyms, see more Miller et al. 1990) connected by a network of morphosemantic relations (Fellbaum et al. 2007). Associated with this is the topological concept of meaning, defining the meaning of a lexical unit through the network of relations in which that unit participates (Maziarz et al. 2013). This approach is embedded in linguistic methodology (see de Saussure 1991:134-135, Lyons 1995:107). These relations occur at the synset and unit both levels. The paper presents, through examples, a comprehensive way of analyzing the meanings of Polish adjectives *pelny* 'full' and pelen 'full', describing relations at both levels, as well as unit attributes such as domain, register, and gloss. From the perspective of semantic syntax, there are two language units (LU) (Bogusławski 1976, Grochowski 1981) based on shape pełny: (1) ktoś/coś jest pełny kogoś/czegoś 'someone/something is full of someone/something' and (2) ktoś/coś jest pełny 'someone/something is full' (Alberski 2016). The existence of these two units is illustrated by the observation of a correct sentence containing both of them simultaneously, see (1) Jej pelne usta sq pelne jedzenia 'her full lips are full of food'. However, from the perspective of natural language processing, greater granularity of meanings is necessary to increase the quality of NLP. The work indicates types of usage expressions of *pelny* 'full' and *pelen* 'full', i.e., connectivity with specific noun classes, e.g., concrete objects having a specified capacity whose volume has reached a very high value: (2) Stary, wypiwszy pełny kubek lury, kiwnął na mnie 'the old man, having drunk a full mug of beer, nodded at me', or abstract objects having a feature that has reached the maximum value, the highest degree of development: (3) Uważają bowiem, że człowiek ma wolność i kontrole nad swoim życiem, za które ponosi pełną odpowiedzialność 'they believe that man has freedom and control over his life, for which he bears full responsibility'. The paper also discusses selected applications of Słowosieci in natural language processing, e.g., automatic definition generation (Wojtasik et al. 2023).

## 11<sup>th</sup> International Conference on Meaning and Knowledge Representation University of Almería, 3-5 July 2024

		No. of lemmas	No. of LUs	No. of synsets	Monosemous lemmas	Polysemous lemmas
	Verbs	20849	45720	34022	11385	9464
	Nouns	135640	180139	135552	110761	24879
plWordNet	Adverbs	8134	14234	11419	4767	3367
	Adjectives	29484	54749	47315	16673	12811
	All	194107	294842	228308	143586	50521
	Verbs	1	1	1	1	0
	Nouns	11133	11296	7514	10980	153
enWordNet	Adverbs	0	0	0	0	0
	Adjectives	1	1	1	1	0
	All	11135	11298	7516	10982	153
	Verbs	11540	25061	13789	6284	5256
	Nouns	117966	146560	82189	102014	15952
Princeton	Adverbs	4475	5592	3625	3729	746
worunet	Adjectives	21808	30072	18185	16876	4932
	All	155789	207285	117788	128903	26886

#### Table 1. http://plwordnet.pwr.wroc.pl/wordnet/stats

#### REFERENCES

- Alberski, Bartłomiej, 2016, "Syntaktyczno-semantyczne właściwości wyrażenia *pełny*", *Różne paradygmaty różne spojrzenia na język*, red. S. Żurowski, Toruń, s. 11-29.
- Bogusławski Andrzej, 1976, "O zasadach rejestracji jednostek języka", Poradnik językowy z. 8, s. 356-364.
- Fellbaum, Christiane, Anne Osherson, and Peter E. Clark, 2009, "Putting semantics into WordNet's" morphosemantic" links." Human Language Technology. Challenges of the Information Society: Third Language and Technology Conference, LTC 2007, Poznan, Poland, October 5-7, 2007, Revised Selected Papers 3. Springer Berlin Heidelberg, 2009.
- Grochowski Maciej, 1982, "Zarys leksykologii i leksykografii. Zagadnienia synchroniczne", Toruń.
- Lyons, John, 1995, "Linguistics Semantics: An Introduction". Cambridge.
- Maziarz, Marek, Maciej Piasecki, and Stanisław Szpakowicz. "The chicken-and-egg problem in wordnet design: synonymy, synsets and constitutive relations." *Language Resources and Evaluation* 47 (2013): 769-796.
- Miller, George A., et al., 1990, "Introduction to WordNet: An on-line lexical database." *International journal of lexicography* 3.4 (1990): 235-244.
- Saussure de Ferdinand, 1991, "Kurs językoznawstwa ogólnego", tł. Kazimierz Polański, Warszawa.
- Wojtasik, Konrad, Arkadiusz Janz, Bartłomiej Alberski and Maciej Piasecki, 2016, "Wordnet for Definition Augmentation with Encoder-Decoder Architecture." Proceedings of the 12th Global Wordnet Conference. 2023, p.50-59.

bartlomiej.alberski@pwr.edu.pl

## METAPHORS IN THE TRENCHES. POLARIZATION IN THE POSTERS OF THE SPANISH CIVIL WAR

Roberto Asenjo Alcalde

Universidad Autónoma de Madrid (Spain)

In 1936 began the Spanish Civil War, an armed conflict which lasted until 1939 with devastating consequences on the Spanish society. Even now the current fracture and polarization in the political spectrum and social fabric seems to be the aftermath of that past belligerent situation in Spain. Apparently, mass media and social networks on the internet have turned into virtual trenches where people attack each other's opinions on diverse issues such as the controversial historic memory law and amnesty law. This fact could indicate that divergent ideologies and opinions are deeply entrenched in the Spanish population's mindset. Apart from the physical violence exerted by both sides in the 1930s struggle, there was an intense war of propaganda on the streets through posters stuck to the walls. This study analyses the metaphorical content of those posters which aimed to influence people's perspective on the conflict. Also this research delves into the interaction of forcedynamic patterns (Talmy, 2000) based on the use of force and other image schemas (Lakoff and Johnson, 1980; Hedblom, 2020) derived from the physical experience of a confrontation. Mainly, the methodology to identify metaphors in these printed pictures is inspired by the Visual Metaphor Identification Procedure (Steen, 2018) and the Target method (Soares da Silva, 2016). Moreover, the most recent discursive approaches to sociopolitical polarization and conflict (Romano, 2024) have been considered. On the whole, the purpose of this research is to shed light on how the persuasive messages conveyed in the posters of the Spanish Civil War became important instruments of propaganda and testimony of a divided society.

### REFERENCES

- Hedblom, M.M. (2020). Image Schemas and Concept Invention: Cognitive, Logical, and Linguistic Investigations. Cognitive Technologies. Springer, Cham. <u>https://doi.org/10.1007/978-3-030-47329-7</u>
- Lakoff, G. & Johnson, M. (1980). Metaphors we live by. University of Chicago Press.
- Preston, P. (2006). The Spanish Civil War: Reaction, Revolution, and Revenge. London, HarperCollins.

Romano, M. (2024) Metaphor in Socio-Political Contexts: Current Crises. Berlin, Boston: De Gruyter Mouton. <u>https://doi.org/10.1515/9783111001364</u>

Soares da Silva, A. (2016). The persuasive (and manipulative) power of metaphor in austerity' discourse A corpus-based analysis of embodied and moral metaphors of austerity in the Portuguese press. In M. Romano & M. D. Porto (Eds.). *Exploring Discourse Strategies in Social and Cognitive Interaction. Multimodal and cross-linguistic perspectives.*, 79-108. Amsterdam/Philadelphia: John Benjamins.

Steen, G. J. (2018). Visual metaphor: Structure and Process. John Benjamins Publishing Company.

Thomas, H. (1979). La Guerra Civil Española. Madrid, Ediciones Urbión.

Talmy, L. (2000). Toward a cognitive semantics. Cambridge (Massachusetts), The MIT Press.

VV. AA. Multiple authors (1981). Carteles de la Guerra Civil Española. Madrid, Ediciones Urbión.

roberto.asenjo@uam.es

## MULTIMODAL METONYMY IN BRANDING: A CORPUS-BASED ACCOUNT OF MOROCCAN FOOD PRODUCTS

Fatima Azzahraa El Yamlahi and María Enriqueta Cortés de los Ríos

Universidad de Almería (Spain)

Armstrong and Kotler (2018) maintain that a brand covers a range of elements, including a name, term, sign, symbol, or design, or a combination of them, intended to identify the goods and services of one seller or group of sellers and to differentiate them from those of competitors. Taking this definition into consideration, whenever a new name, logo or symbol is given to a new product, we can assume that a new brand is created (Keller 2003). Taking into consideration the above mentioned definitions of a brand, we hold that the relation between a brand and a product is metonymic in nature. Metonymy is a figure of thought based on conceptual contiguity (Dirven 2002). This notion of contiguity is used to express "a close or direct relationship between two entities" (Evans and Green 2006: 311). According to this definition, it follows that there is an associative relation between a brand and a product (Azzahraa & Cortés de los Ríos, 2022). Ungerer (2000) maintains that CONTAINER FOR CONTENT and THE NAME FOR PRODUCT are used as effective advertising tools, consciously and extensively. In this regard, conceptual metonymy is likely to play an important role in the meaning construction of brands because usually the product is presented by the brand name, which metonymically stands for the item in question: BRAND NAME FOR THE PRODUCT. To this end, we have selected a corpus of food brands taken from the Office Marocain de la Propriété Industrielle et Commerciale (OMPIC). This work has been structured according to the Cognitive Theory of Metaphor and Metonymy (Lakoff & Johnson 1980, 1999; Ruiz de Mendoza Ibáñez and Galera Masegosa's typology (2014), and building on Peréz Sobrino's applications (2017) of this typology together with Forceville multimodality proposals (1996, 2009). Our aim is to analyse the role of metonymic cognitive operations in these names (domain expansion and domain reduction metonymies together with metonymic chains), and to reveal the modes (monomodal or multimodal) in which the brand names are manifested.

## REFERENCES

- Azzahraa El Yamlahi, F. & Cortés de los Ríos, M <sup>a</sup> E. (2022). Examining the metonymic relation between a brand name and a product: A case study of Moroccan cosmetic brand names. *Ibérica 44*, 391-416
- Armstrong, G. & Kotler, P. (2018). Armstrong, G. and Kotler, P. (2018). *Principles of Marketing*. (17th ed.) Pearson education limited: United Kingdom.
- Dirven, R. (2002). Metonymy and metaphor: Different mental strategies of conceptualization. In R. Dirven & R. Pörings (Eds.). *Metaphor and metonymy in comparison and contrast*. Berlin: Mouton de Gruyter, 75-111.
- Evans, V. & Green, M. (2006). *Cognitive Linguistics: An Introduction*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Forceville, C. (1996). *Pictorial Metaphor in Advertising*. New York: Routledge.
- Forceville, C. (2009). Non-verbal and multimodal metaphor in a cognitivist framework: Agendas for research In C. Forceville and Urios-Aparis (Eds.), Multimodal metaphor (pp. 19-42). Berlin/ New York: Mouton de Gruyter

- Lakoff, G. & Johnson, M. (1980). *Metaphors we Live By*. Chicago: The University of Chicago Press.
- Lakoff, G. & Johnson, M. (1999). *Philosophy in the Flesh: The Embodied Mind and its Challenge* to Western Thought. New York, basic books.
- Keller, K.L. (2003). *Strategic brand Management: Building, Measuring, and Managing Brand Equity.* (2nd ed.). New Jersey: Prentice Hall.
- Pérez Sobrino, P. (2017). *Multimodal metaphor and metonymy in advertising*. John Benjamins Publishing Company.
- Ruiz de Mendoza, F. J., & Masegosa, G. (2014). *Cognitive modeling: A linguistic perspective*. Amsterdam/Philadelphia: John Benjamins.
- Ungerer, F. (2000). Muted metaphors and the activation of metonymies in advertising. In A. Barcelona (Ed.), *Metaphor and metonymy at the Crossroads*, 321–340. Berlin and New York: Mouton de Gruyter.

fatima.azzahraa.elyamlahi@gmail.com; mecortes@ual.es

## COGNITIVE CONDENSATION AND SYNESTHETIC FORCE IN THE WORKS OF SEAMUS HEANEY

## Carmen M. Bretones Callejas

University of Almería (Spain)

This paper explores the concepts of cognitive condensation and synesthetic force in the literary works of Seamus Heaney. By examining Heaney's poems as condensed narratives, the study highlights the narrative and storytelling dimensions present in both poetry and short stories. It investigates the intersections of culture, cognition, and affective resonance that contribute to the lasting impact of his literary craft. Through this lens, the analysis reveals how Heaney's poetry encapsulates complex cognitive and cultural phenomena, reinforcing the iconic and affective power of his storytelling. This perspective shows the symbiotic relationship between cognitive representation and literary form, shedding light on the cognitive and emotional depths that define Heaney's contributions to literature.

cbreton@ual.es

## FREUDIAN METAPHOR: ANALOGIES OF PSYCHOANALYTIC TERMINOLOGY TO SCIENCE, SPACE, AND BODY CONCEPTS

Jana Dietzel

### Stiftung Universität Hildesheim (Germany)

The development of technical language requires specific linguistic features, such as morphological, syntactic, lexical, and text-linguistic characteristics. Moreover, technical language implements a clear linguistic procedure to transmit knowledge in a way that fosters comprehensibility (cf. Peters, 2021: 32). As a young discipline, psychoanalysis needed to establish itself during the early 20th century. It was necessary for the language to balance clarity and precision when addressing the research subjects and methods. A multitude of studies suggest that psychoanalysis draws upon scientific disciplines that predated its founding, which could aid comprehension (e.g. Habermas, 1970; Bally, 1961). Nevertheless, interpretation may also be influenced by cultural, social, political, and epistemological factors of the era when the texts were written, as well as Freud's personal background. As a physician, Sigmund Freud considered himself primarily a natural scientist, despite his research approach diverging from the prevalent somatic methods of his time (cf. Freud 1900: 7).

In several case studies, my dissertation aims to investigate the relationship between scientific terminology and common theories of body and space in the early 20th century. My research focuses on the use of cognitive metaphors in the advancement of psychoanalytic terminology and adopts a qualitative methodology by exploring three psychoanalytic texts authored by Sigmund Freud: "Studies on Hysteria" (1895), "The Interpretation of Dreams" (1900), and "Analysis Terminable and Interminable" (1937). This analysis raises the following research questions: What is the role of metaphors in the formation of psychoanalysis? How did the use of metaphors contribute to the establishment of psychoanalysis as a scientific discipline? How do the metaphors used in these works reflect the historical and cultural context in which they were written? Can a particular consistency or coherence be detected in the metaphorical concepts used?

### REFERENCES

Bally, Gustav (1961). Introduction to Psychoanalysis by Sigmund Freud. Rowohlt: Reinbek.

- Freud, Sigmund and Breuer, Josef (1895). Studies on Hysteria.GW I, 75-312 without Breuer's Contributions. GW supplementary volume.
- Freud, Sigmund (1900). The Interpretation of Dreams. Standard Edition. 4/5.
- Freud, Sigmund (1937). Analysis Terminable and Interminable. GW XVI, 59-99.
- Habermas, Jürgen (1973). Erkenntnis und Interesse. Frankfurt. Suhrkamp Verlag.
- Peters, Joachim (2021). Professional discourse, technical language, and central concepts of palliative care and in the 21st century. A corpus-based linguistic discourse analysis on specialized texts of the years 2000- 2019. FAU Studies from the Faculty of Philosophy Volume 20. Erlangen: FAU University Press.

dietzel@uni-hildesheim.de

## APPLYING CORPUS METHODOLOGY TO DETECT SOCIAL PROBLEMS IN USER-GENERATED CONTENT UNITS

Ángel Felices Lago

Universidad de Granada (Spain)

Rocío Jiménez-Briones

#### Universidad Autónoma de Madrid (Spain)

In the multidisciplinary field of crowdsensing (Li et al., 2017; Wang et al., 2019), aiming to extract insights from social sensors through computational methods (Sakaki et al., 2013; Musto et al., 2015; Arthur et al., 2018), this paper introduces a proposal for identifying social problems pertaining to the domains of Population and Macroeconomics within ALLEGRO (Adaptive muLti-domain sociaL-media sEnsinG fRamewOrk), a smart multimodal system for the prediction of social problems (http://allegro.ucam.edu/). Drawing, among others, upon the works of Periñán-Pascual (2023, 2024a,b), Alameda Hernández (2024), Felices Lago (2024), Jiménez-Briones and Felices Lago (2024) and Ureña Gómez-Moreno (2024), the focus lies on addressing two problematic issues that affect the quality of life of many social media users and that can be the topic of their microtexts in X (formerly Twitter): elderly institutional abuse and black economy. First, the research outlines the conceptual framework adopted for tackling these social problems within DIAPASON (unified hybrId ApProach to microtext Analysis in Social-media dedicated to text cr**O**wdseNsing), which is the ALLEGRO module analysis (http://allegro.ucam.edu/diapason/Index.html). Second, the paper elaborates on the methodology, as well as the challenges, involved in compiling the two specific corpora of tweets related to elderly institutional abuse and black economy. As one of the purposes of the ALLEGRO project is to build a medium-sized corpus of social problems that can serve as a training set for deep learning models in text classification (Periñán-Pascual, 2024b), the proposed corpora will be employed as two of the many specific problem subcorpora that structure the gold standard corpus for the automatic detection of social problems by the ALLEGRO system.

#### REFERENCES

- Alameda Hernández, Ángela (2024) Social media detection of texts on the social problem of violence against women within the multimodal intelligent system ALLEGRO. In Jiménez-Briones, Rocío & Avelino Corral Esteban (eds.) Approaches to Knowledge Representation and Language. Granada: Comares. 63-75.
- Arthur, Rudy, Chris A. BOULTON, Humphrey SHOTTON & Hywell T. P. WILLIAMS (2018). Social sensing of floods in the UK. *PLoS ONE* 13(1). <u>https://doi.org/10.1371/journal.pone.0189327</u>
- Felices Lago, Ángel (2024) Description of social problems by means of schemas related to the Income domain in the DIAPASON platform. In Jiménez-Briones, Rocío & Avelino Corral Esteban (eds.) Approaches to Knowledge Representation and Language. Granada: Comares. 27-43.

- Jiménez-Briones, Rocío & Ángel FELICES LAGO (2024) La formalización del conocimiento en DIAPASON a través de una muestra de problemas poblacionales y macroeconómicos. In Olmo-Cazevieille, Françoise (ed.) *Investigación Lingüística en Entornos Digitales*. Valencia: Tirant Lo Blanch, 187-216.
- Li, Wei, Wen-jun WU, Huai-min WANG, Xue-qi CHENG, Hua-jun CHEN, Zhi-hua ZHOU & Rong DING (2017) Crowd intelligence in AI 2.0 era. *Frontiers of Information Technology* & *Electronic Engineering* 18: 15–43.
- Musto, Cataldo, Giovanni SEMERARO, Pasquale LOPS & Marco de GEMMIS (2015). CrowdPulse: A framework for real-time semantic analysis of social streams. *Information Systems* 54: 127-146.
- Periñán-pascual, Carlos (2023) From Smart City to Smart Society: A quality-of-life ontological model for problem detection from user generated content. *Applied Ontology* 18(3): 263–306.
- (2024a) Modelización de las quejas de los ciudadanos como artefactos digitales culturales: DIAPASON. In Olmo-Cazevieille, Françoise (ed.) *Investigación Lingüística en Entornos Digitales*. Valencia: Tirant Lo Blanch. 129-156.
- (2024b) Exploring problems through social media: The case of Beach Quality. In Jiménez-Briones, Rocío & Avelino Corral Esteban (eds.) Approaches to Knowledge Representation and Language. Granada: Comares. 11-26.
- Sakaki, Takeshi, Makoto OKAZAKI & Yutaka MATSUO (2013). Tweet analysis for real-time event detection and earthquake reporting system development. *IEEE Transactions on Knowledge and Data Engineering* 25(4): 919-931.
- Ureña gómez-moreno, Pedro (2024) Integrating corpus methodology into the construction of an intelligent crowdsensing system. In Jiménez-Briones, Rocío & Avelino Corral Esteban (eds.) *Approaches to Knowledge Representation and Language*. Granada: Comares. 45-62.
- Wang, Dong, Boleslaw K. SZYMANSKI, Tarek ABDELZAHER, Heng JI & Lance KAPLAN (2019) The age of social sensing. *IEEE Computer* 52(1): 36–45.

afelices@ugr.es; rocio.jimenez@uam.es

## GUIDELINES FOR LLM-ASSISTED CORPUS BUILDING AND ANNOTATION

Nicolás José Fernández Martínez

Universidad de Jaén (Spain)

Corpus building and annotation are two time-consuming and labour-intensive tasks that require trained annotators to collect and label texts (Dunn, 2022). With the advent of powerful Large Language Models (LLMs) (e.g., ChatGPT, Bing, Gemini, etc.), a new avenue of potential capabilities is opening up for both tasks. Through prompt engineering, the process of crafting instructions to obtain a desired output (Liu et al., 2023), corpora of different types and for varied purposes can be automatically created and annotated, eliminating the need for manual work. To this end, we present a survey of automatic corpus construction and annotation in the era of LLMs, with a focus on natural language processing (NLP) tasks (Ostyakova et al., 2023; Yu et al., 2023). We also provide several prompting guidelines and examples for both corpus construction and annotation (Törnberg, 2024). Additionally, we demonstrate a practical application by automatically constructing and annotating a synthetic corpus of artificial tweets with negative attitudinal functions (Fernández-Martínez, 2024). The corpus was used to develop and evaluate NLP models in different experiments, showing great accuracy. Our research could pave the way for future LLM-assisted research in corpus building and annotation in computational linguistics.

### REFERENCES

- Dunn, Jonathan. (2022). Natural Language Processing for Corpus Linguistics. Cambridge University Press. <u>https://doi.org/10.1017/9781009070447</u>
- Fernández-Martínez, Nicolás José. (2024). Introducing the NLP task of negative attitudinal function identification. *Procesamiento Del Lenguaje Natural*, 72, 57–71. <u>https://doi.org/10.26342/2024-72-4</u>
- Liu, Pengfei, Yuan, Weizhe, Fu, Jinlan, Jiang, Zhengbao, Hayashi, Hiroaki, & Neubig, Graham. (2023). Pre-train, Prompt, and Predict: A Systematic Survey. ACM Computing Surveys, 55(9). <u>https://doi.org/10.1145/3560815</u>
- Ostyakova, Lidiia, Smilga, Veronika, Petukhova, Kseniia, Molchanova, Maria, & Kornev, Daniel. (2023). ChatGPT vs. Crowdsourcing vs. Experts: Annotating Open-Domain Conversations with Speech Functions. In *Proceedings of the 24th Meeting of the Special Interest Group on Discourse and Dialogue* (pp. 242–254). Association for Computational Linguistics.

Törnberg, Petter. (2024). Best Practices for Text Annotation with Large Language Models, 1–21.

Yu, Danni, Li, Luyang, Su, Hang, & Fuoli, Matteo. (2023). Using LLM-assisted Annotation for Corpus Linguistics A Case Study of Local Grammar Analysis. CoRR, cs.CL/2305. <u>https://doi.org/2010.48550/arXiv.2305.08339</u>

njfernan@ujaen.es

## THE ROLE OF ECHOING IN THE CONVENTIONALIZATION OF IMPLICATIONAL MEANING

Alicia Galera Masegosa

University of Almería (Spain)

Our proposal is framed within the Lexical Constructional Model (LCM), a usage-based cognitively-oriented constructionist approach to language initially proposed by Ruiz de Mendoza and Mairal (2008) and Mairal and Ruiz de Mendoza (2009). One of the proposals of this model is the classification of constructions (understood as fixed form-meaning pairings, cf. Goldberg 1995, 2006) into four levels of meaning representation, namely argument-structure, implicational, illocutionary and discourse constructions.

Non-denotational meaning can be implicational and/or illocutionary in nature. The former captures the speaker's subjective attitude, and the latter involves inferences that arise from sociocultural conventions that regulate social interaction. Frequency of association between implicational meaning and constructional configurations gives rise to implicational constructions. A case in point is Kay and Fillmore's (1999) *What's X Doing Y*? construction, which is often used as a way of expressing the speaker's negative attitude towards X doing Y.

Ruiz de Mendoza & Galera (2014) put forward a set of cognitive operations that are potentially active at all levels of linguistic description. These authors acknowledge the status of echoing (initially proposed by Sperber and Wilson (1986) in their account of irony) as a cognitive operation consisting in the repetition of (part of) a thought, a (real or imaginary) state of affairs or a linguistic expression (Galera, 2020; Ruiz de Mendoza & Galera, 2014).

Along these lines, it is our aim to explore the impact of echoing in the conventionalization of nondenotational meaning. For doing so, we have selected the *Does X Look Like Y*? construction. We suggest that the presence of echoic remarks in the 'Y' slot invariably endows the construction with additional meaning that goes beyond the elicitation of information, which is the primary function of interrogatives. Compare the following:

(1) Do I look like a real front-desk person?

#### (2) Do I look like *I* want to go on a road trip?

Example (1) is uttered in a context in which a person is applying for a job and has dressed up accordingly. The speaker is asking for the hearer's opinion, actually questioning whether he looks like Y. By contrast, the speaker in (2) feels that it is evident that Y is not the case. After the interlocutor has suggested going together to Lake Tahoe for the weekend, the speaker, who is sick in bed, points at his tired face and utters (2). The Y variable of the construction is realized by the echoic mention "I want to go on a road trip", which captures the meaning implications derived from the hearer's proposal. As we shall demonstrate, speakers express an often negative evaluation of some aspects of the hearer's previous discourse and/or behaviour by echoing the hearer's assumptions, which they assume to be evidently incorrect.

Drawing from naturally occurring data (online corpora, movie scripts, etc.), the present study is aimed at analyzing a significant number of instantiations of this construction in order to back up our contentions and provide systematic patterns of analysis that may also be applied to other constructional configurations.

agm011@ual.es

## ENHANCING LINGUISTIC ANNOTATION FOR EMOTION ANALYSIS: A HYBRID METHOD FOR COMPUTER-ASSISTED ANNOTATION

Hongling Gou

Universitat Politècnica de València (Spain)

Natural Language Processing (NLP) plays a significant role within artificial intelligence and linguistics, striving to empower computers with the capacity to comprehend, interpret, and generate human language. NLP techniques are extensively applied in annotation. The purpose of the annotations is to support subsequent applications which can produce some kind of useful result by analyzing texts (Wilcock, 2022). Thus, the annotations of datasets is essential in diverse fields of NLP. Despite its importance, manual annotation is a labor-intensive task, which is timeconsuming and prone to be subjective with human error, hindering the efficiency of emotion analysis and model training. The aim of this study is to investigate if artificial intelligence can help linguists with text annotation. In order to achieve this goal, we designed a hybrid model that combines a quantitative analysis using transformers, qualitative analysis using large language models (LLMs), and manual emotion revision to enhance the annotation of the MBTI Kaggle dataset. This dataset, which was developed in the field of personality detection, contains 8675 user-generated texts collected from the PersonalityCafe forum. According to Fayyad et al. (1996), Knowledge Discovery in Databases methods and techniques are used for making sense of data. Therefore, we applied such methods and techniques, using both manual labeling and automatic labeling with the software TexMiLAB (Periñán-Pascual, 2024), to classify the user-generated texts into five different emotions based on Ekman's (1992) emotion model: joy, anger, fear, surprise, and disgust. Then, we applied the kappa coefficient and LLMs to evaluate and revise the annotation results. Finally, our research findings demonstrate that (i) individual transformers can help linguists to annotate emotions in datasets, (ii) however, the combination of multiple transformers fails to yield enhanced performance in the annotation process, and (iii) in general, computer-assisted techniques have the advantage of assisting human annotation

### REFERENCES

Ekman, P. (1992). An argument for basic emotions. Cognition & emotion, 6(3-4), pp.169-200.

- Fayyad, U., Piatetsky-Shapiro, G., & Smyth, P. (1996). From data mining to knowledge discovery in databases. AI magazine, 17(3), pp.37-37.
- Myers, I. B. (1985). A Guide to the Development and Use of the Myers-Briggs Type Indicator: Manual. Consulting Psychologists Press.
- Periñán-Pascual, C. (2024) *Minería de Textos para Investigadores Lingüistas*, Valencia, Tirant Lo Blanch.

Wilcock, G. (2022). Introduction to linguistic annotation and text analytics. Springer Nature.

hgou@posgrado.upv.es

## PRELIMINARY DESIGN OF AN ONTOLOGY FOR THE DOMAIN OF ORTHOPEDIC DISEASES

Linxi Jiang

#### Universitat Politècnica de València (Spain)

Nowadays, the healthcare industry stands to benefit immensely from natural language understanding systems because of the sheer volume of unstructured medical data being generated, as well as the increasing proliferation of electronic health record (EHR) systems and medical digital devices. Digitalized healthcare data facilitates storage and transmission, but raw data does not meet the requirements for performing mining tasks. Extensive and heterogeneous textual descriptions contained in EHRs pose significant difficulties for the extraction of valuable insights. Utilizing natural language understanding technology to convert unstructured data into a machine-readable form is the foundation of clinical text mining (Hier & Brint, 2020).

Developing intelligent applications based on clinical text mining, such as clinical decision support (Sutton et al., 2020), phenotype extraction (Zhang et al., 2019), computer-aided diagnosis (Wani & Arora, 2020) and disease prediction systems (Shah et al., 2020) is attracting attention of researchers from different domains since it is an interdisciplinary field which involves knowledge, methods and techniques of biomedicine, machine learning, natural language processing (NLP), corpus linguistics, as well as knowledge representation.

As an important semantic technology, previous investigations (Schenk et al., 2023; Xue et al., 2019; Nicolson et al., 2022) have proven the utility of ontological models in knowledge management. In particular, ontologies organize the domain knowledge in the form of relevant concepts and relationships among them, thereby facilitating knowledge integration and discovery (Amith et al., 2018). In the medical domain, ontologies constitute a solid foundation for a variety of healthcare information systems for encoding disease, symptoms, examination results, diagnosis, and treatment in EHRs.

This research focuses on creating an ontology for diseases of the musculoskeletal system and connective tissue using the concepts and entity relationships extracted from the free-text clinical notes database *MIMIC-IV-Note* (Johnson et al., 2023). Furthermore, we plan to incorporate the ontology into an interactive system for queries on orthopedic diseases. In this first phase of investigation, we have reviewed the relevant researches about the application of NLP techniques for medical entity and relationship extraction and the use of ontological models in clinical knowledge engineering. We have created a corpus with 138,832 EHRs extracted from *MIMIC-IV-Note* ensuring that it is processable by computers to complete a series of NLP tasks. In the next phase, we will concentrate on the development of the ontology following METHONTOLOGY (Fernández-López et al., 1997), a methodology to build ontologies from scratch.

As an concrete application of NLP and semantic analysis technologies in the clinical domain, the value of this research is shown in two aspects. On the one hand, it helps to deepen the comprehensive understanding of diseases of the musculoskeletal system and connective tissue. On the other hand, it can improve diagnostic accuracy and increase the overall productivity of the healthcare sector.

#### REFERENCES

- Amith, M., He, Z., Bian, J., Lossio-Ventura, J. A., & Tao, C. (2018). Assessing the practice of biomedical ontology evaluation: Gaps and opportunities. *Journal of biomedical informatics*, 80, 1-13.
- Fernández-López, M., Gómez-Pérez, A., & Juristo, N. (1997). Methontology: from ontological art towards ontological engineering. Proceedings of the AAAI97 Spring Symposium, 33-40.
- Hier, D. B.,& Brint, S. U. (2020). A Neuro-ontology for the neurological examination. *BMC Medical Informatics and Decision Making*, 20, 1-9.
- Johnson, A., Pollard, T., Horng, S., Celi, L. A., & Mark, R. (2023). MIMIC-IV-Note: Deidentified free-text clinical notes (version 2.2). *PhysioNet*.
- Nicholson, N. C., Giusti, F., Bettio, M., Carvalho, R. N., Dimitrova, N., Dyba, T., ... & Martos, C. (2022). A multipurpose TNM stage ontology for cancer registries. *Journal of Biomedical Semantics*, 13(1), 7.
- Schenk, P. M., Wright, A. J., West, R., Hastings, J., Lorencatto, F., Moore, C., ... & Michie, S. (2023). An ontology of mechanisms of action in behaviour change interventions. *Wellcome Open Research*, 8.
- Shah, D., Patel, S., & Bharti, S. K. (2020). Heart disease prediction using machine learning techniques. *SN Computer Science*, 1(6), 345.
- Sutton, R. T., Pincock, D., Baumgart, D. C., Sadowski, D. C., Fedorak, R. N., & Kroeker, K. I. (2020). An overview of clinical decision support systems: benefits, risks, and strategies for success. *NPJ digital medicine*, 3(1), 17.
- Wani, I. M., & Arora, S. (2020). Computer-aided diagnosis systems for osteoporosis detection: a comprehensive survey. *Medical & biological engineering & computing*, 58, 1873-1917.
- Xue, X., Hang, Z., & Tang, Z. (2019). Interactive biomedical ontology matching. *PloS one*, 14(4), e0215147.
- Zhang, Y., Cai, T., Yu, S., Cho, K., Hong, C., Sun, J., ... & Liao, K. P. (2019). High-throughput phenotyping with electronic medical record data using a common semi-supervised approach (PheCAP). *Nature protocols*, 14(12), 3426-3444.

eeevitababy@gmail.com

## PRINCIPLES OF METAPHTONYMY FORMATION (BASED ON ENGLISH VERBS WITH SEMANTICS "TO DIVIDE")

Svetlana Kiseleva and Nella Trofimova

Saint-Petersburg University of Economics (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 metaphor, 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.

svkiseljeva@bk.ru; nelart@mail.ru

## COGNITIVE OPERATION IN CONVERSATIONAL MEANING

## Sanjaya Lenka

## Indian Institute of Technology BHU (India)

Conversational meaning is an inferring meaning about the context and intentionality of a speaker. In this sense, this meaning differs from the literal meaning which is coded in speech. The truth value of this meaning is related to external conditions and the truth value of literal meaning is related to internal conditions. Conversational meaning can express information precisely and concisely which is rarely possible in literal meaning. This meaning can be classified into direct and indirect conversational meaning. The previous type can be contextual meaning, meaning of politeness, meaning of power of social relationship and the later type can be irony, conversational implicature, or suggestive meaning. The important differences between these two are the first is an independent semantic space and the second is an inner product semantic space. Further, the previous one has a certain semantic direction and the latter has no semantic direction, it is no longer an independent semantic space. The source of this meaning is utterance rather than a sentence. Conversational meaning is naturally used in various speech acts, discourses, conversational implicature, and contexts. The important issues of conversational meaning are related to its comprehension and interpretation. It is so complex that the language users cannot easily comprehend it. Specifically, the use of metaphor, irony, and metonymy in discourse and different narratives and knowledge representations create uncertain meanings. The language users cannot easily find the exact information and reality. In natural language processing, a computer never recognizes the conversational meaning of a speech.

This paper tries to develop a theoretical model of conversational meaning that makes it understood and interpretable in speech act, discourse, conversational implicature, and metaphor. It proposes that conversational meaning co-exists with literal meaning and has different dimensions. The literal meaning is the linguistic meaning of utterance and conversational meaning is the contextual and intentional meaning. The important fact is context changes the literal meaning of utterance into a contextual meaning. There is a linear relationship between literal meaning and conversational meaning. Acquiring and understanding this linear relationship and using it in communication is a cognitive operation. This operation can be formalized in logic. This study argues that cognitive operations in conversational meaning can be of two types; decomposition semantic space and inner product semantic space. The decomposition semantic space represents how many conversational meanings are expressed by a literal meaning and how a literal meaning represents many conversational meanings. Decomposition is seen in utterance expressing politeness in imperative utterances, and utterances expressing power of an authoritative relationship.

The second type of cognitive operation can be formalized as inner product semantic space in which conversational meaning is a dependent meaning. It is seen in the expression of irony, metaphor, and conversational implicature. Finally, the paper concludes that understanding cognitive operations can help the audience and computer understand the conversational meaning of discourse, speech acts, and narrations.

### REFERENCES

Austin. J.L. 1955. *How to do things with words*, Oxford, at the Clarendon press. Axler. Sheldon.2010. *Linear Algebra Done Right*, third edition, springer.

- Alan Cruse.D.2000. *Meaning in Language, An introduction to semantics and pragmatics*. Oxford University Press.
- Eco. Umberto. 1976. A theory of semiotics, Indian University Press, Bloomington London.
- Emanuel A. Schegloff. 2017Conversation Analysis in The Oxford Handbook of Pragmatics, Edited by Yan Huang
- Kempson.R.M.1996. Semantic Theory. Cambridge University Press.
- Grice. Paul. 1989. Studies in the Way of Words, Harvard University Press.
- Grundy. Peter. 1999. Doing pragmatics. Black well
- Matilal. B.K.1992. The Word and the World, Oxford University Press

slenka.hss@itbhu.ac.in

## METAPHORS AND THE CONSTRUCTION OF CLIMATE CHANGE DISCOURSE. A CORPUS BASED STUDY OF MULTIMODAL METAPHORS

Sayf Mohamed

Paris Cergy University (France)

Metaphor is a cognitive tool that operates in both verbal and non-verbal communication, or in various modes simultaneously (Forceville, 2010). This paper explores the interrelation of conceptual, and visual metaphor in the online discourse on climate change. To this aim, the analysis integrates the Conceptual Metaphor Theory by Lakoff and Johnson (1980) and the Appraisal theory of evaluation by Martin and White (2005) to analyse conceptual and visual metaphors representing climate change in France. Following a multimodal approach to metaphor, posts by the French president Macron on the social media platform X in the year 2023 were collected and annotated for conceptual metaphors and for Attitudinal evaluation. The results reveal a network of metaphors which describe the issue as a journey, a battle or a natural disaster. These representations contribute to legitimate France's Energy Sobriety plan. This study contributes to the research on metaphorical evaluation by integrating conceptual metaphor analysis and evaluation in online discourse.

### REFERENCES

Charteris-Black J. 2004. Corpus approaches to critical metaphor analysis. Hamsphire.

- Flusberg, S. J., Matlock, T., & Thibodeau, P. H. (2018). War metaphors in public discourse. *Metaphor and Symbol*, 33(1), 1–18. <u>https://doi.org/10.1080/10926488.2018.1407992</u>
- Gibbs, R. W. (1994). *The poetics of mind: Figurative thought, language, and understanding.* Cambridge University Press.
- Hood, S., and Martin, J. (2005). Invoking attitude: the play of graduation in appraising discourse. Revista Signos, 38(58), 195 - 220.
- Jungherr, A. (2016). Twitter in politics: A comprehensive literature review. SSRN Electronic Journal. <u>https://doi.org/10.2139/ssrn.2402443</u>
- Kövecses, Z. (2013). Metaphor and emotion. Cambridge University Press.
- Lakoff, G., & Johnson, M. (1980). Metaphors We Live by. Chicago: University of Chicago Press.
- Lakoff, R. T. (1990). Talking power: The politics of language in our lives. Basic Books.
- Martin, J., & White, P. (2007). The language of evaluation. Basingstoke: Palgrave Macmillan.
- Thibodeau, P. H., & Boroditsky, L. (2011). Metaphors we think with: The role of metaphor in reasoning. *PLoS ONE*, 6(2). <u>https://doi.org/10.1371/journal.pone.0016782</u>

seyfmac@hotmail.com

## LEXICAL BLENDS IN SPANISH: AN ANALYSIS FROM THE POINT OF VIEW OF COGNITIVE LINGUISTICS

Carla Ovejas Ramírez and M.ª Sandra Peña Cervel

University of La Rioja (Spain)

Globalization and new technologies have blurred the boundaries between different languages. This is one of the reasons why word formation processes that were not common in a particular language are becoming key processes for the creation of neologisms in that language. Word formation is one of the main areas of study within lexicology and lexicography (Algeo 1977, Bauer 1983, Plag 2003, Harley 2006, Gramley 2012). Some of the main functions of creating new lexemes are filling certain lexical gaps or simply endowing language with creativity. One of these word formation techniques that has traditionally not been considered popular but is gaining ground in Spanish is blending. A lexical blend results from the fusion of two or more lexical units, at least one of which has undergone a shortening that does not respect regular morphological segmentation.

Lexical blends are the second most frequent word formation process in English, preceded by composition (Peña 2022). To carry out our study, we have compiled a corpus of 200 lexical blends in Spanish taken from the latest updates of new lexemes (especially those created between 2014 and 2019) from the Banco de Neologismos of Centro Virtual Cervantes. Following the work by Peña (2022), where she analyzes English lexical blends from the point of view of cognitive modeling, in this proposal, we will provide a classification of lexical blends in the Spanish language from our corpus based on (1) the semantic type of lexical blend: endocentric (one of the base words functions as the semantic head and the other base word as the modifier. They are hyponyms of one of their members (e.g., sexilio)), exocentric (neither of the base words is the semantic head. They are not hyponyms of any of their constituents (e.g., *redneck*)), or dvandva (both base words can be considered the nucleus of the lexical blend (e.g., glocal)); (2) the morphosyntactic patterns of each lexical blend (e.g., in dramedia, the morphosyntactic pattern is composed of two nouns); and (3) the cognitive operations underlying the creation of each lexical blend. A cognitive operation is a mental mechanism aimed at constructing a semantic representation from linguistic input to make it meaningful in context (Ruiz de Mendoza & Galera 2014). Cognitive operations have been classified as formal and content cognitive operations. The latter are low-level mechanisms that play a crucial role in making inferences in the process of constructing meaning. The analysis of our data shows that parameterization underlies the conceptual configuration of many of our examples (e.g., in basuraleza, the term "basura" is parameterized as a specific type of trash found in nature; therefore, this is a case of parameterization based on location). We will analyze how this and other cognitive operations underlie the creation of different lexical blends in our corpus in order to compare our study of lexical blends in Spanish with that carried out by Peña (2022) of English lexical blends.

### REFERENCES

Algeo, J. 1977. Blends, a structural and systemic view. *American Speech 52*, 47–64.
Bauer, L. 1983. *English Word Formation*. Cambridge: Cambridge University Press.
Gramley, S. 2012. *The History of English: An Introduction*. USA/Canada: Routledge.
Harley, H. 2006. *English Words: A Linguistic Introduction*. MA/Oxford: Blackwell Publishing.

Peña, M.S. 2022. Lexical blending in terms of cognitive modeling. In *Figurativity and Human Ecology*. Editado por Alexandra Bagasheva, Bozhil Hristov y Nelly Tincheva. [Figurative Thought and Language series, 17, 2022, pp. 275–304].

Plag, I. 2003. Word-Formation in English. Cambridge: Cambridge University Press.

Ruiz de Mendoza, F. J., & Galera, A. (2014). *Cognitive modeling. A linguistic perspective*. Amsterdam/Philadelphia: John Benjamins.

carla.ovejas@unirioja.es; sandra.pena@unirioja.es

## KNOWLEDGE GRAPHS INTERPRETATION FOR THE SPEECH AND LANGUAGE ANALYSIS OF A MULTIMODAL AI AGENT FOR PRE-SCREENING COGNITIVE FUNCTION OF OLDER ADULTS

#### Kulvinder Panesar

#### University of Bradford (United Kingdom)

### María Beatriz Pérez Cabello de Alba

#### Universidad Nacional de Educación a Distancia (Spain)

Cognitive dysfunction is demonstrated by a decline in cognitive abilities such as language, visuospatial skills, memory, judgment, and mental agility. Each affected person will experience symptoms differently. Over the past decade, there has been an increase in people challenged by cognitive decline due to several factors such as health, lifestyle, ageing, employment, and mental health.

This work is part of a wider project which aims to create a computerised pre-screening tool to assess cognitive function with a main focus on language processing with speech and language analysis as a health assessment of cognitively healthy older adult participants. The tool will take an individual through a set of tasks which will assess their ability to produce and understand language, to plan, organize, and recall information as per the task and session. This wider project spans computer science, artificial intelligence (AI), neuropsychology, language, healthcare, and user experience. This work builds on previous research [1-3] where a model was created and trained using the Dementia TalkBank [4]. This project has an existing conceptual architecture with an RRG [5] language engine with interfacing phase models (language, cognitive and dialogue) to provide the means of understanding the speech utterance and performing linguistic and grammatical analysis [6]. It also has ongoing development work using automatic speech recognition (ASR) and creation of an AI multimodal assessment platform.

Our focus here is the language component after the ASR to text and RRG grammatical analysis stages. Next, the qualitative assessment of the utterances is conducted forming sets of small trees of language production parameters (lexical syntax, semantic, discourse and pragmatic), and language cognition and speech parameters (ontology, lexicon, speech) for each utterance of each task of each participant session, forming a complex network of the qualitative assessment data. This data, represented as vector embeddings, is exported to form a domain in a labelled property knowledge graph (KG) using the technology environment Neo4j[7]. Here, we create a knowledge model of interlinked descriptions of concepts, entities, and relationships whereby nodes are the entities, and the links are relationships between the nodes. Further the KGs will be enriched with domain (constraint) rules on the structure and content of the graph to maintain data integrity, and indexes to improve performance. Invoking Neo4j's Cyber (declarative query language) queries in Neo4j will enable vector similarity search/ pattern matching (of similar results - similar or repeated cognitive behaviour), un-directed transversals (for example - the relationship between tasks and utterances) and aggregations (totals). The outcome of this will be a set of interpretations of the language production and language cognition and speech ability, that recognises normal and abnormal data patterns. For the wider project it will help to identify at-risk individuals and the early detection of cognitive decline coupled with the quantitative and multimodal (and visuospatial) assessment.

## REFERENCES

- [1] K. Panesar and M.-B. Pérez-Cabello de Alba. "RRG 2021 Program Entry Abstract: Analysing language and cognition of patients living with dementia: a view from a RRG perspective." York University. <u>https://rrg2021.apps01.yorku.ca/program/</u> (accessed 4 Aug 2021.
- [2] K. Panesar and M. B. Pérez Cabello de Alba, "Natural language processing-driven framework for the early detection of language and cognitive decline," *Elsevier - Language and Health*, vol. 1, no. 2, pp. 20-35, 2023/10/04/ 2023, doi: https://doi.org/10.1016/j.laheal.2023.09.002
- [3] K. Panesar and M.-B. Pérez-Cabello de Alba, "A hybrid conceptual model using statistical and linguistic methods for prescreening memory through language and cognition," in *Approaches to Knowledge Representation and Language*, R. Jiménez-Briones and A. Corral Esteban Eds. Granada: Editorial Comares, 2024, ch. In Press, pp. 111-134.
- [4] A. M. Lanzi *et al.*, "DementiaBank: Theoretical rationale, protocol, and illustrative analyses," *American Journal of Speech-Language Pathology*, vol. 32, no. 2, pp. 426-438, 2023.
- [5] R. D. Van Valin Jr, *Exploring the syntax-semantics interface*. Cambridge: Cambridge Univ Press, 2005.
- [6] K. Panesar, "An Evaluation Of A Linguistically Motivated Conversational Software Agent Framework," *Journal of Computer-Assisted Linguistic Research*, vol. 3, no. 3, pp. 41-66, 2019, doi: <u>https://doi.org/10.4995/jclr.2019.11118</u>
- [7] Neo4j, "Neo4j Graph Technology," 2024. [Online]. Available: <u>https://neo4j.com/</u>

bperez-cabello@flog.uned.es; k.panesar@bradford.ac.uk

## THE CONSTRUCTIONAL NATURE OF MERISM

M.ª Sandra Peña Cervel and Carla Ovejas Ramírez

University of La Rioja (Spain)

Merism has been studied as a metonymy-related figure of speech (Ruiz de Mendoza, 2020; Peña, 2022; Peña and Ruiz de Mendoza, 2022) within the framework of Cognitive Linguistics. It has been traditionally defined as a particular kind of synecdoche. Peña (2022) has provided an exhaustive study of this figure making use of some of the tools of cognitive modeling. From this perspective, merism is a combination of two or three words which share the same word class, whose function is to make reference to a whole (e.g., *flesh and bone*, "the body"), and in which the relevant subdomains designed by the two or three lexical units grant access to a whole conceptual domain. Thus, expansion is a key ingredient in the characterization of merism. This figure is usually a fixed combination. The role of contrast in merism has been one of the major concerns in the study of merism. However, Peña (2022) found that contrast is not a sine qua non of this figure. On the basis of this observation, this author has distinguished between contrastbased merism (e.g., rich and poor, "everyone") and bare merism, which is not based on contrast (e.g., collar and tie, "formal clothing"). Metaphor and metonymy, the master tropes, being conceptual phenomena, are very difficult to identify and retrieve from a corpus in spite of the great efforts made in this direction over the past thirty years, in which corpus-based methods have been adopted as the dominant linguistic approach (Stefanowitsch, 2006, p. 1). However, merism is always constructionally cued. This contribution addresses the constructional nature of merism from the point of view of cognitive modeling. A preliminary attempt has been made by Peña (2022: 245-247). This proposal goes beyond this initial analysis by providing an in-depth study of meristic constructions which are in need of further investigation (e.g., 'X and Y alike', as in "young and old alike") or ('X as well as', as in "small as well as large") from the point of view of cognitive modeling, to this end, a corpus of 500 examples has been gathered from different sources (Corpus of Contemporary American English, Google, online dictionaries (e.g., *Cambridge Dictionary Online*), and the traditional literature on merism).

### REFERENCES

- Peña, M.S. 2022. For better, for worse, for richer, for poorer, in sickness and in health: A Cognitive-Linguistic approach to merism. *Metaphor and Symbol* 37(3), 229–251.
- Peña, M.S., & Ruiz Ruiz de Mendoza, F.J. 2022. Figuring out Figuration. A Cognitive Linguistic Account. Amsterdam/Philadelphia: John Benjamins.
- Ruiz de Mendoza, F.J. 2020. Understanding figures of speech: Dependency relations and organizational patterns. *Language & Communication* 71, 16–38.
- Stefanowitsch, A. 2006. Corpus-based approaches to metaphor and metonymy. In Corpus-based approaches to metaphor and metonymy. A. Stefanowitsch & S.T. Gries (eds.). Berlin/New York: Mouton de Gruyter, pp. 1–16.

sandra.pena@unirioja.es; carla.ovejas@unirioja.es

## SYNTHETIC DATASETS FOR COMMUNITY PROBLEM DETECTION IN ALLEGRO

Carlos Periñán-Pascual

Universitat Politècnica de València (Spain)

Text generation with pre-trained large language models (LLMs) has accelerated the progress in natural language generation, which can produce high-quality texts in tasks like summarization, creative writing, and question answering, among others. In this regard, customizing the generated content and style involves imposing task-specific constraints through prompts, which consist of natural-language instructions and examples provided to the LLM to guide its output. One application of controllable text generation using LLMs is the creation of synthetic corpora. In this study, we present the process of developing synthetic corpora within the ALLEGRO project framework, a multimodal system aimed at understanding society through social media content. ALLEGRO (Periñán-Pascual, 2023) analyzes user-generated content to extract relevant knowledge for identifying community problems. Among its modules, DIAPASON examines English and Spanish text messages to detect community problems using natural language processing, text mining, and knowledge engineering techniques. In the DIAPASON ontology, each community problem type is defined by a problem schema, a formal representation containing semantic concepts and pragmatic functions to address the meaning dimensions of each problem type. The challenge lies in employing these problem schemas to automatically generate prompts that guide the creation of a robust synthetic training dataset. Such a dataset should enable DIAPASON to classify real-world texts covering various community problem types.

### REFERENCES

Periñán-Pascual, Carlos (2023) "From Smart City to Smart Society: A quality-of-life ontological model for problem detection from user-generated content". Applied Ontology 18 (3), pp. 263-306.

jopepas3@upv.es

## FUNCTIONAL CONCEPTS OR FUNCTIONS AND CONCEPTS? BOUNDARIES OF THE CONCEPTUAL-PROCEDURAL DISTINCTION AND LSA

Adrián Raso González

Complutense University of Madrid (Spain)

Latent semantic analysis is a computational tool for describing the acquisition and representation of knowledge in natural language (Landauer and Dumais, 1997). Since its first applications in the 1980s, its value has been revealed in its ability to detect (1) semantic dependencies and (2) textual patterns (Landauer, 2006) across the large document masses on which the analysis is applied. As a device we can find its use especially in the natural language processing literature and at present its methodology has influenced outstanding technologies such as transformers, which take advantage of the quantification in weights that LSA started by gathering in its matrices. In this paper we wish to approach the central problem of procedural semantics, that is, to find a sharp criterion for distinguishing between procedural meanings (namely, those that we traditionally conceive of as encoded by functional categories, such as prepositions or determiners) and conceptual meanings (which manifest some form of reference, and classically underlie lexical categories) (Wilson, 2011). We believe that extracting information through LSA can yield interesting properties of the expressions that make up these classes, not only by looking at their proportion of occurrence in texts, but also by understanding what distributions they obey and what relationships terms of one genre of meaning and another may develop at the latent level.

We believe that functional and conceptual expressions are organised in semantic spaces in the sense of Landauer and Dumais, 1997, obeying a distribution in essence similar to that which follows from the distinction between lexical layer and functional layer in generative grammar. The semantic spaces in this key would be composed of a network of dependencies at the latent level, which would be controlled by the argument structure as well as by ontological information such as can be found in models like Pustejovsky, 1995. From LSA, we would be importing seminal notions like *lag* to measure linear distributions in sentences (Howard Bing Jing, Addis and Kahana, 2006), among others. This would make it possible to make explicit in such a predictively effective vocabulary both the nature of both groups of entities in natural language, and their behaviour especially in semantic and syntactic matters.

We will consider the corpus British Academic Written English (BAWE) with, on the one hand, one hundred terms from different specific domains which will be mostly nouns, so that they will carry a conceptual character, and on the other hand, the terms which make up the inventory of procedural categories in English, from prepositions and determiners to connectors. We have chosen BAWE because its collection of texts aimed at specific purposes guarantees a consistency in vocabulary that facilitates the construction of the types that will constitute the matrix.

We seek to detect dependencies between terms of conceptual and functional nature in order to describe at a latent level the structures served by both entities. If that is not given in a reasonable way, we would try to address how is the relation between these two forms of meaning.

### REFERENCES

Bresnan, J., Asudeh, A., Toivonev, I. & Wechsler, S. (2016). *Lexical-Functional Syntax*. Wiley Blackwell.

- Connell, L. M. (2000). Categories, Concepts and Co-ocurrence: Modelling Categorisation Effects with LSA. University of Edinburgh.
- Curcó, C. (2011). On the Status of Procedural Meaning in Natural Language. In V. Escandell-Vidal, M. Leonetti & A. Ahern (Eds.), *Procedural Meaning: Problems and Perspectives*.
- Howard Bing Jing, M. W., Addis, K. M. & Kahana, M. J. (2006). Semantic Structure and Episodic Memory. In T. Landauer, D. McNamara, S. Dennis & W. Kintsch (Eds.) Handboof of Latent Semantic Analysis.
- Escandell-Vidal, V. & Leonetti, M. (2011). On the Rigidity of Procedural Meaning. In V. Escandell-Vidal, M. Leonetti & A. Ahern (Eds.), *Procedural Meaning: Problems and Perspectives*.
- Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological Review*, 104(2), 211–240.
- Landauer, T. K., Foltz, P. W., & Laham, D. (1998). An introduction to latent semantic analysis. *Discourse Processes*, 25(2–3), 259–284.
- Louwerse, M. M. (2006). Symbolic or Embodied Representations: A Case for Symbol Interdependency. In T. Landauer, D. McNamara, S. Dennis & W. Kintsch (Eds.) *Handboof of Latent Semantic Analysis*.
- Pustejovski, J. (1995). The Generative Lexicon. The MIT Press.
- Kintsch, W. (2006). Meaning in context. In T. Landauer, D. McNamara, S. Dennis & W. Kintsch (Eds.) *Handboof of Latent Semantic Analysis*.
- Wilson, D. (2011). The Conceptual-Procedural Distinction: Past, Present and Future. In V. Escandell-Vidal, M. Leonetti & A. Ahern (Eds.), *Procedural Meaning: Problems and Perspectives*.

adrraso@ucm.es

## TWO LEVELS OF LEXICAL AMBIGUITY AND A UNIFIED CATEGORICAL APPROACH

Chenchen Song

Zhejiang University (China)

The formal representation of meaning plays a key role in both theoretical and computational linguistics, and there is a gap between lexical and compositional semantics in both fields. To attain a more human-like understanding of meaning, this division of labor must eventually be bridged. One class of solutionsuses Category Theory as the bridge, a main representative of which is the DisCoCat framework<sup>[1]</sup>. It linkslexical and compositional semantics via a functor between a logical syntax category and a vector space semantics category. However, due to the particular syntactic and semantic environments used in this framework, it is difficult to transfer linguists' accumulated research results into it.

In this paper, I explore a different categorical approach to semantics (building on [2]), which has a more practical interface between theoretical and computational linguistics. Specifically, I demonstrate how two levels of lexical ambiguity can be integrated into meaning representation and how this expanded linguistic view better reflects the nuanced nature of human language.

The two levels of lexical ambiguity I investigate are respectively at the categorization and the post-categorization level. In current theoretical linguistics, a lexical word can be decomposed into a functional *categorizer* and an idiosyncratic *root*<sup>[3]</sup>. Thus, *fish* (n.) is given the subatomic representation [N  $n \sqrt{\text{FISH}}$ ].At this level, ambiguity arises when the same root combines with different categorizers (e.g., [V  $v \sqrt{\text{FISH}}$ ]).Categorization-level ambiguity is also observed in "semilexical" words<sup>[4]</sup>, as exemplified in (1).

(1) a. [N  $n \sqrt{BA}$ ] 'handle', [V  $v \sqrt{BA}$ ] 'to hold' [Mandarin Chinese]

b. [Cl Cl  $\sqrt{BA}$ ] 'a classifier for holdable objects', [P  $p \sqrt{BA}$ ] 'a preposition marking direct objects'

NB *ba* still has lexical idiosyncrasies in its functional uses. Thus, the prepositional *ba* sounds neutral andis replaced by *jiang/gei* in more elevated/casual contexts. In [2], the categorizer-root combination is given acategorical semantics via a writer monad<sup>[5]</sup>: the categorizer is interpreted as a pure function, while theroot, by a "side effect," as in (2).

(2) a.  $\llbracket[N n \sqrt{BA}]\rrbracket = write(\sqrt{BA}) >>= \lambda y.\eta(\llbracket n \rrbracket) = \langle \llbracket n \rrbracket, \{\sqrt{BA}\}\rangle$  (an entity idiosyncratically characterized by  $\sqrt{BA}$ )

b.  $\llbracket [C1 \ C1 \ \sqrt{BA}] \rrbracket = write(\sqrt{BA}) >>= \lambda y.\eta(\llbracket C1 \rrbracket) = \langle \llbracket C1 \rrbracket, \{\sqrt{BA}\} \rangle$ 

(a classifier denotation idiosyncratically characterized by  $\sqrt{BA}$ )

The same categorical semantics has been extended to emojis and similar multimodal elements in [6].

Post-categorization ambiguity arises after the categorization step, such as the ambiguity of *book* (n.)between its physical reading and its informational reading, as in (3).

(3) a. The book is interesting. (informational reading)

b. The book weighs five pounds. (physical reading) ([7], p. 86)

Such ambiguity can be modeled via a categorical treatment of Pustejovskyan "dot types," as in  $(4)^{[8]}$ .

(4) a. *book* :::  $P \cdot I \equiv BOOK \rightarrow P \times I$  (the noun *book* has type  $P \cdot I$ , aka BOOK, which is a subtype of  $P \times I$ )

**b** 
$$P \times I \xrightarrow{\pi_1} P \rightarrow R$$
,  $P \times I \xrightarrow{\pi_2} I \rightarrow E$  (basic types P and I can be retrieved, which are subtypes of E)

The categorical environment in this paper is a topos, so we can straightforwardly add "dot types" toit and model post-categorization ambiguity alongside categorization ambiguity. I further show that we can conveniently represent the intuitive connection between the two levels of ambiguity, as in (5).

!

This is a usual subobject-classifying pullback in a topos. Our rationale is that "dot types" are assigned towords out of context, whereas denotations like  $\langle [n], \{ \forall BOOK \} \rangle$  are of specific word instances in concrete contexts. Hence, the type of  $\langle [n], \{ \forall BOOK \} \rangle$  is an already disambiguated "dot aspect" (via (4b)).

The two levels of lexical ambiguity above are not only fundamental to human language but also at the crossroads of (morpho)syntactic and semantic analyses. Our categorical approach provides a practical channel for research results from theoretical linguistics to be transferred to NLU system building.

#### REFERENCES

- [1] Coecke, B., M. Sadrzadeh, and S. Clark. 2010. Mathematical foundations for a compositional distributional model of meaning. *Linguistic Analysis* 36(1–4), 345–383.
- [2] Song, C. 2021. On the semantics of root syntax: Challenges and directions. *Proceedings of the 18th Logic and Engineering of Natural Language Semantics workshop (LENLS18)*, 61–74.
- [3] Marantz, A. 1997. No escape from syntax: Don't try morphological analysis in the privacy of your own lexicon. In A.Dimitriadis et al. (Eds.), *University of Pennsylvania Working Papers in Linguistics*, Vol. 4.2, 201–225.
- [4] Song, C. 2019. On the formal flexibility of syntactic categories. University of Cambridge dissertation.
- [5] Asudeh, A. and G. Giorgolo. 2020. *Enriched meanings: Natural language semantics with Category Theory*. Oxford:Oxford University Press.
- [6] Song, C. 2022. Sentence-final particle vs. sentence-final emoji: The syntax-pragmatics interface in the era of CMC. Talkat SyntaxLab, University of Cambridge, June 28.
- [7] Asher, N. 2011. *Lexical meaning in context: A web of words*. Cambridge: Cambridge University Press.
- Banonnaud, W. 2019. A topos-based approach to building language ontologies. In R. Bernardi,
   G. Kobele, and S. Pogodalla (Eds.), *Formal Grammar: 24th International Conference, FG* 2019, Riga, Latvia, August 11, 2019, Proceedings, 18–34. Berlin: Springer.

cjs021@zju.edu.cn

## TRANSLATING CULTURAL KNOWLEDGE REPRESENTATION FOR HERITAGE PRESERVATION

Hamest Tamrazyan

Ukrainian State Dragomanov University (Ukraine)

Gayane Hovhannisyan

Brusov State University, Yerevan, Armenia

The paper discusses key conceptual aspects of Digital Humanities as a paradigm aiming at cultural knowledge representation and heritage preservation for the organization, accessibility, and interoperability of diverse heritage collections by translating them into a common conceptual system of description. Given the need for adjusting diverse cognitive systems, the intercultural translator acts as a cognitive linguist whose task is to compare the conceptual structures, detect and translate the elements of cultural knowledge with high precision and clarity of the information structure.

To solve the problems of transmitting culture-specific knowledge into a universal system of representation researchers acknowledge the need for an interdisciplinary approach to developing metadata for artifacts of diverse cultures, besides the field-specific vocabularies, ontologies, taxonomies, and terminology: deep knowledge of history, comparative culture studies, etymology and historical linguistics, literature ethnography. The artifacts belong not only to cultural sites and objects but also to the linguistic consciousness shaped over centuries in the practice and communication of the community. The common Linked Open Data for both professional and popular science communication gains universal value only if it reflects and precisely incorporates these semantic data into the Simple Knowledge Organization System.

The study claims that developing knowledge organization systems is not just a technical procedure but a scrupulous documentation process with deep language and culture research in the background. KO systems are influenced by social structures and power dynamics and language change. In this respect, linguistic exteriorization of subjective, culture-specific knowledge that is meant for internal uses is the core of SKOS frames and categorization. Therefore, KO system development involves factors and criteria ensuring inclusivity and multiple facets of reflecting the objective reality.

One of the factors mentioned above is the historical, social and cultural context of knowledge informing the KO systems with critical evaluation adaptability in the light of changing societal norms and values. The systemic context approach is especially pivotal in our days of hybrid wars where cultural heritage can be distorted and weaponized for political manipulation. We advocate for those standards and methods of KOS, to disarm such manipulation by proliferating diversity and inclusivity.

Our approach ensures scientific objectivity in KO systems design by employing critical examination of power structures and discourse strategies within which they operate knowledge representation. This method is highly important while working with cultural and historical materials, where broad factual knowledge and a deep understanding of historical, cultural, and political contexts are essential.

Examples of preserving the endangered Armenian epigraphy in Artsakh and Ukraine illustrate the key statements introduced in the paper along with the new tools we developed to adapt the existing ontologies and taxonomies to our task of a single cultural knowledge translation and representation from different sites and contexts.

### REFERENCES

- Tamrazyan, H.; Hovhannisyan, G. Digital Guardianship: Innovative Strategies in Preserving Armenian's Epigraphic Legacy. *Heritage* 2024, 7, 2296-2312. <u>https://doi.org/10.3390/heritage7050109</u>
- Tamrazyan, H. (2023). "Digitization of the Inscriptions on the Monuments of Armenian Cultural Heritage in Nagorno-Karabakh Region". Publish in - Baillot, Anne, Tasovac, Toma, Scholger, Walter, & Vogeler, Georg. (2023, July 1). Digital Humanities 2023: Book of Abstracts. Digital Humanities 2023. Collaboration as Opportunity. (DH2023), Graz, Austria. <u>https://doi.org/10.5281/zenodo.7961822</u>
- Lavios, S. (2014). Corpus-based Translation studies: Research and Application. London: Bloomsbury Academic.
- Orlandi, Sabtucci, Casarosa, R.V., Liuzzo, P.M. (2014). Information technologies for epigraphic and cultural heritage. The first EAGLE international conference.

hamest.tamrazyan@epfl.ch; grhovhannisyan@bryusov.am

## CONTEXT THEORY VERSUS POLITENESS THEORY APPLIED TO THE GREAT FORTUNE AND SPOILT CITY NOVELS BY OLIVIA MANNING TO REFUTE ROMANIAN CRITICS AND CENSORS

#### Cristina Zimbroianu

#### Polytechnic University of Madrid (Madrid)

Olivia Manning (1908-1980) was an English writer best known for a series of six novels she wrote after the Second World War, which were published as two trilogies: The Balkans Trilogy (1981) and The Levant Trilogy (1982), both based on her experiences during the war. The novels The Great Fortune (1960) and The Spoilt City (1062), which are part of the Balkan trilogy, are set in Bucharest, Romania, where the couple Harriet (Olivia Manning) and Guy Pringle (Reggie Smith, her husband) lived for a year, as Guy had a lecturing post at the University of Bucharest. Manning recounts the historical events of 1939-1940 and how they affected Romanian society and the English community in Bucharest. However, her experiences in Romania were not well received by Romanian critics and censors. They considered Manning's allusions to the poor peasants, the beggars, the English's patronising view of Romania and the situation of the Jews to be inaccurate, and also accused her of being an untalented, uncultured writer, a snob, trying to describe the people of Bucharest without racist undertones. This article therefore aims to analyse Manning's discourse in order to clarify whether the opinions of critics and censors are accurate or not. So far, no further research has been conducted in this direction, as the articles published on Manning's trilogies have focused on the reception of The Great Fortune by the Romanian censors, the perception of Romania by the British and the evaluation of some of Manning's comments by literature students (Zimbroianu 2021, Goșa, 2019 and Macari 2016).

Therefore, this present analysis is supported by the fact that Manning's trilogies are based on her experiences during the Second World War. Thus, she conveys to the reader the reality she encountered in Romania in the years 1939-1940. And so this study is divided into two parts: Firstly, Manning's discourse is analysed from the perspective of Brown and Levinson's (1987) politeness theory in *Politeness Some Universals in Language Usage* and secondly, from the perspective of Teun van Dijk's (2008) context theory as presented in *Discourse and Context. A Sociocognitive Approach*. The results show that when politeness theory is applied, in cases such as the one in which Manning considers the peasants to be less than beasts because "*they had not the beauty or dignity of beasts. They treated their animals and their women with the simple brutality of savages*", then there is a clear threat to the peasants' face. On the other hand, if we consider the context of the speaker, in this case Manning, and in this context the 'mental model' introduced by Dijk (2008 and 2009), which is subjective and involves personal representations of an event, then Manning's view that the peasants are less than beasts is created by her own emotions, perspective, judgement and subjective perception of the situation which requires an in depth study that would justify Manning's comments.

#### REFERENCES

Brown, P., & Levinson, S. C. (1987). *Politeness: Some universals in language usage*. Cambridge University Press.

Dijk T. van. (2008). Discourse and Context: A Sociocognitive Approach. Cambridge University Press.

- --- (2009). Society and Discourse: How Social Contexts Influence Text and Talk. Cambridge University Press.
- Manning, O (1960). The Great Fortune. William Heinemann.
- --- (1962). The Spoilt City. William Heinemann.
- --- (1981). The Balkan Trilogy. Penguin Books.
- --- (1982). The Levant Trilogy. Penguin Books.
- Goșa, C. (2019). "Through a Reader Looking Glass. Olivia Manning's *The Balkan Trilogy*". *Romanian Journal of English Studies*. University of Timișoara: Timișoara, pp. 54-64.
- Macari, I. O. (2016). "British Travellers Stick Together Olivia Manning's Gulliver in the Balkans". Acta Iassyensia Comparationis, 18 (2/2016), Călătorii și călători / Voyages et voyageurs / Travels and Travellers. University Alexandru Ioan Cuza: Iași, pp. 89-96.
- Zimbroianu, C. (2021). "The Reception of Olivia Manning's *The Great Fortune* in Romania". *Connections: A Journal of Language, Media and Culture*. University of Alberta: Edmonton Canada, pp. 18-29.

cristina.zimbroianu @upm.es