

La 10^{ème} édition du colloque : INFOL@NGUES 2018

L'Informatique au service des Langues

19, 20 et 21 avril 2018

Béja, Tunisie

Préface

La sélection d'articles publiés dans le présent recueil constitue les actes de la 10^{ème} édition du colloque INFOLANGUES 2018 qui s'est déroulé à Béja du 19 au 21 avril 2018.

Le colloque INFOLANGUES ambitionne comme à l'accoutumée de regrouper dans un même lieu scientifique, chercheurs, informaticiens et linguistes, pour échanger des idées et des approches dans le domaine du traitement des langues. Donc c'est un défi de rassembler une panoplie de chercheurs travaillant sur plusieurs thèmes abordés dans ce colloque, tels que Linguistique computationnelle, Linguistique pour la reconnaissance de l'écriture, Langue et informatique, Dictionnaires numériques, Fouille de données textuelles, Traduction automatique (TA) et automatisée (TAO), Ontologies, Evaluation des modèles de langues par les TICs, Enseignement virtuel, Recherche d'information.

Le colloque faut il le souligner attire de plus en plus des chercheurs nationaux et internationaux. Dans cette édition nous avons accepté 35 papiers rédigés par des informaticiens et des linguistes, en trois langues Arabe, Français et Anglais. Nous avons invité plusieurs Professeurs et Docteurs, spécialistes dans le domaine du Traitement des Langues Naturelles (TALN), pour présenter l'état d'avancement de leurs recherches au cours des journées du colloque. Nous avons programmé un atelier animé par le professeur Mustafa Jarrar de l'Université Birzeit sur des logiciels sur la traduction automatique.

Remerciement

Nos vifs remerciement vont tous d'abord aux auteurs pour leurs communications, aux relecteurs, membres du comité scientifique, dont les rapports d'évaluation circonstanciés et constructifs ont contribué à améliorer significativement la qualité des articles.

Nos remerciements vont également à tous les invités du colloque qui ont donné le plus par leurs implications et présence. Nous citons entre autre le Prof. Mustafa Jarrar, le Prof. Amor Lehcen et le Dr. Faiza Derbel.

Nous remercions aussi toute l'équipe du comité d'organisation pour leur travail efficace et leur mobilisation continue.

Enfin nous remercions spécialement pour leur soutien financier et aides, l'université de Jendouba, l'Institut Supérieur des Langues Appliquées et d'Informatique de Béja, le Laboratoire de recherche en Technologies de l'Information et de la Communication & Génie Electrique à l'ENSIT, l'Association Actes Académiques et l'usine Kromberg & Shubert.

Comités du colloque

Infol@ngues 2018 est organisé par l'Institut Supérieur des Langues Appliquées et d'informatique de Béja en collaboration avec les deux Laboratoires de recherches LaTICE, Université de Tunis et LLTA, Université de Sfax.

Comité Scientifique

Président :

- **Prof. Rim Faiz**, Université de Carthage, Tunisie.

Membres :

- **Prof. Mourad Elloumi**, Université d'El Manar, Tunisie.
- **Prof. Sami Faiz**, Université de la Manouba, Tunisie
- **Prof. Mohamed Gamoudi**, Université de la Manouba, Tunisie.
- **Prof. Faiez Gargouri**, Université de Sfax, Tunisie
- **Prof. Mohamed Jabir**, Université de Carthage, Tunisie.
- **Prof. Mohamed Jemni**, Université de Tunis, Tunisie.
- **Prof. Ridha Mami**, Université de la Manouba, Tunisie.
- **Prof. Yahya Slimani**, Université de la Manouba, Tunisie.
- **Prof. Mohamed Bouattour**, Université de Sfax, Tunisie.
- **Prof. Ezzedine Zagrouba**, Université d'El Manar, Tunisie.
- **Prof. Eric Wehrli**, Université de Genève, Suisse.
- **M.C. Lamia Hadrich Belguith**, Université de Sfax, Tunisie.
- **M.C. Akila Baklouti**, Université de Sfax, Tunisie
- **M.C. Tahar Labbassi**, Université de Tunis, Tunisie.
- **M.C. Ricco Rakotomalala**, Université de Lyon, France.
- **Dr. Imed Ben Amar**, Université de la Manouba, Tunisie.
- **Dr. Afef Kacem**, Université de Tunis, Tunisie, Tunisie.
- **Dr. Ramzi Farhat**, Université de Tunis, Tunisie.
- **Dr. Bilel Gargouri**, Université de Sfax, Tunisie.
- **Dr. Salma Jammoussi**, Université de Sfax, Tunisie.
- **Dr. Sahbi Hidri**, Université de Tunis, Tunisie.
- **Dr. Karim Chine**, Université de Cambridge, Royaume-Uni.
- **Dr. Abdessatar Mahfoudhi**, Université de Kuwait, Kuwait.
- **Dr. Slim Mesfar**, Université de Jendouba, Tunisie.
- **Dr. Othman Talbi**, Université de la Manouba, Tunisie.
- **Dr. Ghassan Mourad**, Université Libanaise, Liban.
- **Dr. Aymen Khlifi**, ENS Lyon, France.
- **Dr. Talel Zid**, Université de Jendouba, Tunisie.
- **Dr. Asma Bouhafs**, Université de Carthage, Tunisie.
- **Dr. Imed Mahnane**, Université de Jendouba, Tunisie.
- **Dr. Sofien Haboubi**, Université de Tunis, Tunisie.
- **Dr. Faiza Derbel**, Université de la Manouba, Tunisie.

Table de matières

Colloque invités

| | |
|---|---|
| <i>The Arabic Ontology and Lexical Semantic Resources,</i> Mustafa Jarrar | 1 |
| <i>من تطبيقات اللسانيات الحاسوبية : المعاجم الإلكترونية</i> Lahcen Amor | 2 |
| <i>Creating Your Own Blended Course,</i> Faiza Derbel | 3 |

Apprentissage Automatique du Langage

| | |
|---|---|
| <i>Apprentissage de la langue Arabe en utilisant des outils issus du TAL,</i> AbdelKarim Mars, Wafa Touihri & Mohamed Amine Harabi | 4 |
| <i>Reconnaissance des caractères arabes manuscrits par le réseau de neurones profond,</i> Wafa Touihri, AbdelKarim Mars & Mohamed Amine Hrab i..... | 5 |
| <i>Apprentissage tout au long de la vie via un PLE assisté,</i> Belhassen Guettat & Ramzi Farhat | 6 |
| <i>المعلم و مهامه في البيئة التعليمية الحاسوبية،</i> Nadia Hasnaoui | 7 |
| <i>Patterns of Reading Errors among Second Grade Emirati Students,</i> AbdelAziz M. Sartawi | 8 |
| <i>La génétique éditoriale à l'épreuve des humanités numériques : exemple de deux logiciels d'alignement de texte,</i> Ludivine Rey | 9 |

Traitement de la sémantique

| | |
|---|----|
| <i>La complexité sémantique à travers la polysémie: une véritable problématique pour la traduction,</i> Ifaoui Imen | 10 |
| <i>La temporalité : modélisation de la structure syntaxique et interprétation automatique du sens,</i> Siwar Mabrouk | 11 |
| <i>Combinaison «insérer» avec « sous », « sur » : description sémantico-syntaxique en vue de traitement automatique,</i> Randa Kalamoun | 12 |

| | |
|--|----|
| <i>Traitement automatique de la polysémie : le cas des connecteurs dans la phrase complexe,</i> Sahar Sokraf | 13 |
| <i>Approche sémantico-interactive pour un logiciel de traduction juridico-automatique français / tunisien,</i> Tawfik khil | 14 |
| <i>L'apport du traitement du signal de parole aux études phonétiques en sciences du langage,</i> Sabrina Mhathbi | 15 |
| <i>Traitement automatique et analyse sémantique du mot « sécurité » dans les Tweets d'Emmanuel Macron et Marine Le Pen,</i> Chaabani Hayet | 16 |

Traduction Automatique

| | |
|---|----|
| <i>Les verbes savoir/ penser/ croire : problème de traduction automatique,</i> Rym Salem | 17 |
| <i>Les limites de la traduction automatique et les solutions envisagées,</i> Malek Khaldi | 18 |
| <i>Traduction financière et nouvelles technologies de l'information et de la communication (NTIC) : Cas de la « monétique » : moyens de paiement électronique,</i> Leila Laimeche | 19 |
| <i>Les difficultés et les limites de la traduction automatique de la métaphore dans le discours politique (Français / Arabe),</i> Marwa Ammar | 20 |

Traitement du discours

| | |
|---|----|
| <i>Mind Mapping in Classroom: Inspire and Aspire with Coggle,</i> Majoul Boutheina | 21 |
| <i>Tools for Technology-Enhanced Formative Assessment,</i> Asma Gaaloul | 22 |
| <i>L'image et son discours : analyse pragmatique d'une photo,</i> Hajer Mzoughi | 23 |
| <i>L'apport du traitement automatique l'analyse du discours : mesure de l'objectivité du discours journalistique,</i> Jamil Ghouaidia | 24 |
| <i>Le discours électronique extrémiste,</i> Yosra Sayeh | 25 |

| | |
|--|----|
| <i>La polysémie du marqueur évidentiel «évidemment» dans le discours de Chirac,</i> Adel Mnefgui | 26 |
|--|----|

Inférence

| | |
|---|----|
| <i>Les inférences valides vs les inférences valables,</i> Mohamed Bouattour | 27 |
|---|----|

| | |
|---|----|
| <i>Dictionnaire électronique de la prédication épistémique : le nom,</i> Abdessattar Bouzid | 28 |
|---|----|

| | |
|---|----|
| <i>L'inférence automatique (le moteur de recherche : Google) et l'inférence humaine,</i> Abir Berfifa | 29 |
|---|----|

| | |
|--|----|
| <i>Défigements à La Une de la presse écrite tunisienne,</i> Amine Bou ghammouira | 30 |
|--|----|

| | |
|--|----|
| <i>La prédication du SN déterminé par l'adjectif démonstratif dans le cadre anaphorique,</i> Sana Abdelhamid | 31 |
|--|----|

Analyse de sentiments

| | |
|--|----|
| <i>Emotions et interactivité sur Twitter,</i> Najeh Elouni | 32 |
|--|----|

| | |
|--|----|
| <i>Description et méthodologie de traitement d'un corpus de tweets : cas de pragmatèmes d'affect,</i> Najwa Gharbi | 33 |
|--|----|

| | |
|--|----|
| <i>Etude des collocations verbales de sentiment en français et en arabe : à visées applicatives en traitement automatisé du lexique de sentiment,</i> Safa Zouaidi | 34 |
|--|----|

| | |
|---|----|
| <i>الاسم والفعل في التعبير عن الألم: دراسة من خلال آلية المسح العرفني،</i> Habib Mkadmini | 35 |
|---|----|

| | |
|--|----|
| <i>Relations publiques, langue et technologie,</i> Sally Hammoud , | 36 |
|--|----|

| | |
|---|----|
| <i>Problématique des signes de ponctuation en traduction (humaine, TA, TAO),</i> Ghassan Mourad | 37 |
|---|----|

The Arabic Ontology and Lexical Semantic Resources

Mustafa Jarrar
Université Birzeit, Palestine
mustafajarrar@gmail.com

Abstract: The Arabic Ontology and Lexical Semantic Resources. The importance of linguistic ontologies, wordnets, and lexical semantic resources is increasing for building modern and multilingual applications, such as e-commerce, biomedical informatics, information search and retrieval, question-answering and NLP-based tasks, sentiment analysis, interoperability and data integration, big data, among others. This talk overviews the Arabic Ontology, its upper levels and design principles; as well as the digitization of 150 Arabic multilingual lexicons.

من تطبيقات اللسانيات الحاسوبية : المعاجم الإلكترونية

أ.د. عمر لحسن

قسم اللغة العربية وآدابها، جامعة باجي مختار - عنابة (الجزائر)

lahcenamor@yahoo.fr

الملخص: إن شك في أهمية اللغة ضرورية لأسسها الإلكترونية، وذلك بالنظر إلى الأهمية التي تكتسبها اللغة في حياة هذا الإنسان، فهي وسيلة تضمن تواصله مع باقي أفراد مجتمعه، وهي كذلك وسيلة للتعبير عن فكره، وعقيدته، وتجاربها، وعلومه المختلفة، فلا يستطيع إنسان أن يستغني عنها أبداً. فظهرت على مر العصور نظريات متنوعة عادت إلى دراسة الألسن من جميع مستوياتها الصوتية والصرفية والنحوية والدلالية والمعجمية، التي انتهت بها المطاف خلال القرن الماضي إلى بلورتها في إطار نظري وإجرائي حديث يعنى إلى إعطائها طابعاً علمياً من "اللسانيات".

ومن ظهورها في بداية القرن العشرين، ما فتئت اللسانيات تبحث عن أفضل المناهج والنظريات لدراسة اللغة كعلمية، وما فتئت تتفرع إلى فروع شتى وتتقاطع مع علوم كثيرة، فظهرت مجالات بحثية جديدة وعلوم بينية، مثل اللسانيات النفسية psycholinguistique و اللسانيات الاجتماعية sociolinguistique، و اللسانيات الرياضية linguistique mathématique، و اللسانيات الحاسوبية linguistique informatique.

وقد عت اللسانيات الحاسوبية، نذ ظهورها إلى حلّ المشكلات اللسانية والإلكترونية التي كانت اللسانيات التقليدية قد عجزت عن معالجتها، فولجت بإيديين كثيرة، في محاولة لفرض فسها على اللسانيات التقليدية، لتصبح الأبحاث في مجال اللسانيات الحاسوبية هي المسيطرة على البحث اللساني، بما فتحت من مجالات بحث جديدة، وبما قد تهيئ من حلول لمشكلات لسانية وغير لسانية، وبما هلت به البحث للبلاتيين.

ومن بين المجالات التي ولجتها اللسانيات الحاسوبية، ما عرف في مجال البحث اللساني بالتطبيقات التطبيقية، مثل الترجمة الآلية (أو الترجمة بمساعدة الحاسوب TAO) وتعليمية اللغات بمساعدة الحاسوب (E-dico)، الذكاء الاصطناعي intelligence artificielle، المعاجم الإلكترونية E-dico.

وحاول في هذه المداخلة حلّ ودراسة المجالات السابقة، وهو مجال المعاجم الإلكترونية في اللغة العربية، تتبعين أنواعها (المعاجم الألفية اللغة، المعاجم الثابتة، المعاجم في الإنترنت، المعاجم العربية المحملة في الأقراص المدمجة) وما تقدّمه من حلول للبلاتيين والمتعلمين، والمشكلات اللسانية التي تطرأها.

Creating Your Own Blended Course

Faiza Derbel

Université de la Manouba

fderbel26@gmail.com

In this workshop I will introduce the concept and models of Blended Learning and then give some problem-solving activity for the participants to decide what to include in the blend and how to deliver the course.

Apprentissage de la langue Arabe en utilisant des outils issus du TAL

Abdelkarim Mars, Mohamed Amine Harabi, Wafa Touihri

Laboratoire S&T, TIME Higher School, Ariana, Tunisie

Abdelkarim.Mars@gmail.com, wafa.touihri@time.ens.tn et mohamedamine.harabi@time.ens.tn

Abstract : Cet article traite l'application et l'intégration des technologies du traitement automatique de langue (TAL) dans l'apprentissage des langues assisté par ordinateur (ALAO). Nous rapportons notre recherche sur les plates-formes d'éducation assistée par ordinateur. Certaines des méthodes décrites dans cet article ont déjà pris part aux applications de l'utilisateur final qui sont en usage quotidien, d'autres encore attendre pour leur mise en œuvre sous la forme de produits logiciels. Notre but principal dans ce papier est de montrer que la technologie TAL, même dans la forme imparfaite de l'état de l'art actuel, peut accroître aujourd'hui et d'une manière significative l'apprentissage et l'enseignement de langue assisté par ordinateur. Comme une étude de cas, ce document démontre l'application des idées données pour l'apprentissage de la langue arabe.

Keywords: Apprentissage, assisté par ordinateur, plates-formes, Arabe.

Reconnaissance des caractères arabes manuscrits par le réseau de neurones profond

Wafa Touihri, Abdelkarim Mars, Mohamed Amine Harabi

Laboratoire S&T, TIME Higher School, Ariana, Tunisie

wafa.touihri@time.ens.tn, mohamedamine.harabi@time.ens.tn

et abdelkarim.mars@gmail.com

Résumé : Hormis sa nature cursive, le script arabe présente un schéma linguistique complexe qui pose de nombreuses problématiques pour les systèmes automatiques de reconnaissance de l'écriture manuscrite. Dans la conception d'un système de reconnaissance optique de l'écriture manuscrite (OCR), d'un côté, la segmentation des mots reste un travail difficile et qui nécessite beaucoup de temps et de calcul, d'un autre côté, les informations locales sont quelque peu négligées dans les systèmes basés sur une analyse globale qui peut réduire considérablement les performances.

Pour remédier à ces problèmes, des approches de réseau neuronal profond ont été proposées pour la reconnaissance des caractères arabes manuscrits. Ces tentatives de recherche restent limitées et ne prétendent pas satisfaire toutes les complexités de la reconnaissance automatique de l'écriture manuscrite.

Mots clés : reconnaissance; OCR; réseau de neurone, profonds.

Apprentissage tout au long de la vie via un PLE assisté

Belhassen GUETTAT, Ramzi FARHAT, Mohamed JEMNI

Laboratoire de recherche en Technologies de l'Information et de la Communication & Génie

électrique, Université de Tunis 5 Av. Taha Hussein, Bab Menara, Tunis, Tunisie

belhassen.guettat@wanadoo.tn, ramzi.farhat@esstt.rnu.tn, mohamed.jemni@alecso.org.tn

Résumé : Le concept d'environnement d'apprentissage personnel (en anglais PLE pour Personal Learning Environment) suscite ces dernières années un grand intérêt pour le potentiel qu'il offre pour supporter l'apprentissage. Les PLE sont des environnements dont les outils offerts et les contenus qu'ils délivrent sont choisis par l'apprenant. Ceci va lui permettre d'avoir une formation sur mesure et personnelle. Toutefois, la mise en place d'un tel environnement demande des compétences et des efforts de la part de l'apprenant qui ne sont pas garantis. Ceci présente un défi majeur qui laisse poser des questions sur la réalisabilité du concept PLE. C'est dans la lumière de ce défi que nous avons développé un PLE construit autour d'une approche d'assistance. Le long de ce papier, nous allons présenter les composants de ce PLE assisté. De plus, nous allons souligner son adéquation pour l'apprentissage tout au long de la vie. Pour démontrer la faisabilité de notre approche, nous avons développé un prototype, ensuite, nous l'avons expérimenté pour évaluer l'efficacité du processus de l'apprentissage dans un PLE assisté.

Mots clés : PLE assisté, Lifelong Learning, Approche d'assistance.

المعلم و مهامه في البيئة التعليمية الحاسوبية

نادية سناوي

جامعة باجي مختار، عنابة الجزائر

nadia_hasnaoui23@hotmail.com

المخلص: إنَّ الدَّورَ الذي يؤديه المعلم في البيئة التعليمية دورٌ مهمٌ للغاية باعتبارَه أنَّه أحد الأركان الأساسيات للعملية التعليمية لأنَّه يعمل على تقديم مادةٍ ذاتِ قيمةٍ وفقاً لطريقةٍ معينةٍ في بيئٍ صقلٍ تعلم ذو مستوى تعليميٍ متميزٍ - وهذا لا يصبو إليه - و لكونه المتعلم يركز العملية التعليمية يجب مراعاة بوليه و تعداداته عن بناء المادة الدراسية وكذلك يجب مراعاة جوانب أخرى عند تقديم المادة الدراسية من قبل الأتاذنها الغزو التكنولوجي الذي يشهده المجتمع لأنَّه في اللغة العربية أصبحت ملة في ظل التلايد و ينفرونها و هذا لطريقة النمطية التي ينتهجها الأتاذ في تقديم المادة التعليمية فكونه تعلمي اليوم تأثيرين و واكبين للتطور التكنولوجي - الحاسوب - الذي أصبح تلاماً لدى جلمهم ففي البيئة التعليمية العادية يعتمد فيها الأتاذ المادة على الويلة التعليمية العادية - السبورة - فتلقى عليه جميع الأدوات فيلجأ إلى الطريقة التقليدية في تقديم المادة مع أنَّ المقابلة بالكفاءات تشرك التلميذ في بناء الدس، إلا أنَّ الضعف القاعدي الذي يعاينيه التلاميذ خال ذلك و لهذا عمل على توظيف التقنيات التعليمية الحديثة التي تعد خة من الروخ التي تستند عليها دة المستقبل.

Patterns of Reading Errors among Second Grade Emirati Students

Abedalaziz Mustafa Abdelmoti Sartawi

Special Education Department, College of Education, UAE University

asartawi@uaeu.ac.ae

Abstract: The impact of phonological awareness on reading skills is the lead subject of study in the field of literacy, while morphological awareness has been neglected, although it, too, is highly related to reading proficiency. The aim of this follow-up study was to examine second-grade students' reading error patterns by mapping a threefold paradigm (i.e., difficulty level, lexical category, and reading error patterns). The study sample included 168 Emirati second grade students (87 male and 81 female), taken from the Al Ain Educational Zone in the United Arab Emirates. In addition, 144 words were randomly selected from the Arabic reading curriculum. Those words were classified into lexical categories as nouns (45%), verbs (34%), adjectives and adverbs (10%), and function words (i.e., preposition and conjunctions; 11%). Next, the words were classified into three levels of linguistic difficulty (i.e., easy, difficult, and very difficult) according to their morphological complexity. The results revealed that reading *difficulties may be* the result of a threefold interrelated paradigm: the difficulty level of the words (i.e., morphological complexity), lexical category (i.e., nouns, verbs, adjective and adverbs, and prepositions and conjunctions), and pattern of reading errors (e.g., omission of a letter or a syllable and reading the geminated letter as ingeminated). The results also indicated that identifying these paradigms in each student can be useful in early detection of reading difficulty and thus providing an *appropriate remedial program* and determining the emphasis and strategies of instructional programs.

La génétique éditoriale à l'épreuve des humanités numériques : exemple de deux logiciels d'alignement de texte

Ludivine Rey

Université Paris-Sorbonne Lettres

rey_ludivine@yahoo.fr

Résumé : Je vous propose de présenter un exemple de génétique textuel à travers l'utilisation de deux logiciels de traitement automatique des langues : Versioning machine et Médite. Ces logiciels me permettent d'étudier et d'analyser les caractéristiques lexicales, syntaxiques et stylistiques des différentes versions des pièces de théâtre de Pierre Corneille. Ce dramaturge a revu et corrigé ses premières pièces de théâtre à onze reprises pour essayer d'atteindre la forme la plus parfaite que peut avoir une comédie. Par ailleurs, le goût et les attentes des spectateurs changent, évoluent entre l'écriture de ses premières pièces dans les années 1630 et l'édition de 1644, ou bien celle de 1660. Ce sont les deux éditions qui ont reçu le plus de corrections de la part de leur auteur. De plus, on ne s'exprime pas de la même manière en 1633 et en 1682. Notre corpus se limitera donc aux premières œuvres de Pierre Corneille. La première version sera l'édition originale de chaque pièce de théâtre, et elle sera confrontée à l'édition *in-folio* de 1663. Nous étudierons les variantes de deux états des pièces de théâtre de Pierre Corneille en fonction de quatre critères : l'insertion, la suppression, le remplacement et le déplacement. Ensuite, nous les associerons aux catégories syntaxiques et sémantiques des mots ou groupes de mots pour donner naissance à des opérations linguistiques de réécriture. Ainsi, nous terminerons cette analyse par une interprétation génétique de passages choisis.

La complexité sémantique à travers la polysémie: Une véritable problématique pour la traduction

Ifaoui Imen

Laboratoire de Recherche LLTA, Faculté des Lettres et des Sciences Humaines de Sfax

Imen_ifaa@live.fr

Résumé : Le sens linguistique a toujours fait l'objet d'étude des linguistes au sein du même système linguistique, et il constitue un problème fondamental dans le domaine de la traduction. Parmi les problèmes liés au sens, nous avons choisi de travailler sur *la polysémie*, caractère d'un signe qui possède plusieurs sens et qui change de sens selon le contexte et d'un énoncé à un autre. Le traitement de la polysémie met en place des difficultés majeures qui se rencontrent par les différents systèmes de traduction comme celui de *Google Translate* surtout lorsque le transfert se fait entre deux langues, deux systèmes différents dans leur fonctionnement et qui n'ont aucun lien tels que l'arabe et le français.

Identifier le sens de l'unité polysémique n'est pas une tâche commode non seulement pour le récepteur humain mais cela peut offenser le sens global quand il s'agit d'une *traduction automatique* à cause de plusieurs difficultés qui se mettent en place comme le problème de *fidélité*, *l'interférence*, *l'unité minimale* de la langue qui sert de point de départ de la traduction, les *niveaux de traduction*, *les transformations* dans la traduction.

Mots clés : polysémie, traduction, traduction automatique, Google Translate, problème de traduction.

La temporalité : modélisation de la structure syntaxique et interprétation automatique du sens

Siwar Mabrouk

Université de Gafsa, Laboratoire LLTA, Université de Sfax

siwar.mabrouk@yahoo.fr

Résumé : Le traitement automatique des propositions subordonnées temporelles introduites par les conjonctions et les locutions conjonctives implique certainement une approche modulaire, et ce, en élaborant une analyse morphologique de ce type de mots subordonnants afin d'analyser le sens des phrases qui les contiennent. Pour ce faire, nous étudierons essentiellement l'interaction qui se trouve entre les trois modules (morphologique, syntaxique, sémantique) et nous expliciterons la manière par laquelle ces trois modules interviennent dans le traitement automatique.

Notre intervention s'articulera autour de quatre points. Nous expliquons tour à tour l'analyse morphologique et le traitement automatique de la forme, l'analyse syntaxique et la modélisation de la phrase entière ainsi que l'étude sémantique et l'interprétation automatique du sens.

Combinaison «insérer» avec « sous », « sur » : description sémantico-syntaxique en vue de traitement automatique

Randa Kalamoun

La Faculté des Lettres et Sciences Humaines de Sfax, Laboratoire Langage et Traitement
Automatique (LLTA), Université de Sfax
randakalamoun@yahoo.fr

Résumé : Dans cette contribution, nous nous proposerons d'examiner le fonctionnement sémantico-syntaxique du verbe « insérer » qui se combine avec la préposition « sous » et « sur » dans une perspective contrastive français/ arabe.

La variété du sens de ce même verbe engendre une difficulté au niveau de la traduction.

Nous allons étudier dans cet article le processus d'ambiguïté de la traduction automatique du verbe «insérer».

Nous procéderons à la traduction français/arabe d'un corpus phrastique pour essayer d'appréhender les convergences et les divergences qui subsistent dans ces deux systèmes linguistiques. Lors de la traduction, il faut prendre en considération les spécificités des deux langues. Néanmoins, les résultats donnés par la traduction automatique sont peu crédibles. Il fallait l'intervention du cerveau humain pour trancher et éviter l'ambiguïté.

Dans un premier volet, nous utiliserons Antconc, Tropes et les verbes français des Dubois pour faire l'extraction d'un corpus phrastique contenant des exemples du verbe « insérer ». Les particularités sémantiques de ce verbe rendent compte de son comportement syntaxique.

Et dans un second volet, nous étudierons le processus de la traduction automatique de ce verbe ainsi que l'ambiguïté qui en découle.

Mots clés : verbe, préposition, traduction, français/arabe, traitement automatique

Traitement automatique de la polysémie :

le cas des connecteurs dans la phrase complexe

Sokraf Sahar

Laboratoire Langage et Traitement Automatique (LLTA), Université de Sfax
sokraf.sahar@yahoo.com

Résumé : Il s'agit d'examiner une propriété fondamentale des Langues Naturelles et qui constitue autant d'obstacles au traitement automatique. En fait, nous nous sommes intéressés à la polysémie-source de prolifération des ambiguïtés qui peuvent se présenter sur différents niveaux : lexical, sémantique et logico-pragmatique. L'accent a été essentiellement mis sur le traitement automatique de la polysémie des prédicats de second ordre dans la phrase complexe, appelée aussi par Gross (1970) la connexion transphrastique. Celle-ci renferme une trace de codage qui n'est pas uniforme. A noter, en outre, que ces prédicats/connecteurs sont des marqueurs de relation transphrastique exprimant la cause, la conséquence, le temps, le but... Ils sont également des unités non seulement monolexicales présentant un codage adéquat, transparent et explicite (*parce que, à cause de, dans le but de...*) mais aussi des unités polylexicales introduisant ce qu'on appelle surcodage (*dans l'espoir de...*) et sous codage (les temporelles, par exemple, ne codent pas le temps mais une autre relation logique qui est atteinte par inférence). Nous avons, par la suite, vérifié si la résolution de la polysémie s'appuie sur le contenu de noyau prédicatif de P1/P2 tout en nous fondant sur la Théorie des Blocs Sémantiques de Carel (1995). A signaler que l'extraction des phrases en question est effectuée à partir du logiciel *Antconc*.

Mots clés : polysémie, codage adéquat, sous codage, surcodage, prédicat de second ordre, inférence, connexion transphrastique, implicite, Théorie des Blocs sémantiques

**Approche sémantico-interactive
pour un logiciel de traduction juridico-automatique français / tunisien.**

KHIL Tawfik

Faculté des Lettres et Sciences Humaines, Laboratoire du langage et traitement automatique.

L.L.T.A, Université de Sfax

Laboratoire, le Centre Interlangues Texte Image Langage (EA4182), Université de Bourgogne

khiltawfik@yahoo.fr

Résumé : Le présent article est une exposition des principes cruciaux pour une traduction juridico-automatique admise de l'arabe vers le français. Par des enchaînements théoriques et pratiques, nous assisterons à une démonstration d'une traduction juridico-automatique postulant la double performance ; la qualité linguistique et la sécurité juridique. Pour y arriver, trois étapes seront proposées. La perspective méthodologique de ces trois étapes sera le traitement du terme juridique d'abord, dans la langue source et dans le système juridique du départ. Ensuite, son traitement dans le transfert vers la langue cible et le système juridique d'arrivée.

En effet, notre objectif est d'installer un mécanisme technique, dit un logiciel informatique, pour traduire les articles juridiques de la langue française vers la langue arabe. Certes, ce projet de traducteur juridique exige deux postulats. D'abord, la qualité linguistique et, ensuite, la sécurité juridique.

L'apport du traitement du signal de parole aux études phonétiques en sciences du langage

Mhedhbi Sabrine

Institut Supérieur des Langues de Tunis, Université de Carthage

sabina_tunisia@live.fr

Résumé : Le Traitement Automatique des Langues est un domaine pluridisciplinaire se situant à la frontière de l'informatique et de la linguistique. En effet, l'automatisation des données langagières n'a vu le jour qu'avec le développement de systèmes informatiques qui nous a offert la possibilité de la manipulation du langage humain dans tous ses aspects. Parmi les domaines qui se sont basés sur les logiciels dans le traitement et l'analyse des activités langagières, on cite la phonétique qui ne s'est développée que grâce au progrès technique. Dans ce cadre, je mentionne que mon travail de thèse, appartenant au domaine indiqué ci-dessus, vise essentiellement à étudier les structures accentuelles du français par des locuteurs natifs de l'arabe tunisien à travers une étude expérimentale fondée sur la notion de corpus. Les données de ce corpus ont été numérisées directement sur un PC à l'aide du logiciel de traitement du signal PRAAT. Ce dernier a permis le traitement direct des niveaux physiques des productions et la mesure de leurs corrélats acoustiques : la fréquence fondamentale (F0), les analyses spectrales, la durée et l'intensité. A cela s'ajoutent d'autres fonctionnalités qu'on découvrira au fur et à mesure de la lecture de cet article.

Mots clefs : accent, corpus, informatique, linguistique, logiciel PRAAT.

Traitement automatique et analyse sémantique du mot « sécurité » dans les Tweets d'Emmanuel Macron et Marine Le Pen

Chaabani Hayet

Faculté des Lettres et Sciences Humaines de Sfax, LLTA, Université de Sfax

chaabanihayet441@yahoo.com

Résumé : Cet article contribue à montrer la complémentarité entre la linguistique et l'informatique à travers l'utilisation des outils informatique (logiciel, plate-forme) dans l'extraction et l'analyse d'un corpus numérique. Nous avons focalisé notre étude sur le mot « sécurité » dans les tweets de deux comptes Twitter de deux candidats à la présidentielle « Marine Le Pen » et « Emmanuel Macron », lors de leur campagne électorale 2017. Notre objectif est de voir comment elle se représente la question de sécurité chez les deux candidats et par conséquent quels programmes ont-ils préparés pour assurer la sécurité dans le pays ?

Mots-clés : sécurité, Emmanuel Macron, Marine Le Pen, Tropes, IDEO2017

Les verbes *savoir* et *connaître* : problème de traduction automatique

Rym SALEM
Laboratoire LLTA, Université de Sfax
rymme.salem@gmail.com

Résumé : L'objectif de notre travail consiste à examiner les convergences et les divergences dans l'emploi des verbes *savoir* et *connaître* lors du passage de la langue source (français) vers la langue cible (arabe). Notre moteur de traduction automatique est Google Translate. Nous aurons le soin ainsi d'établir une comparaison entre traduction automatique et traduction humaine.

Les limites de la traduction automatique et les solutions envisagées,

Malek Khaldi

Laboratoire Langage et Traitement Automatique (LLTA), Université de Sfax

khaldimalek@gmail.com

Résumé : Notre travail porte sur l'étude des verbes se construisant alternativement avec « à », « de » et parfois avec un infinitif pur. L'usage variant des prépositions qui suivent ces verbes se trouve expliqué en psychomécanique du langage (G. Guillaume) par l'impact du sémantisme des prépositions sur le verbe recteur et sur l'infinitif. En se combinant avec l'une de ces prépositions, le verbe dénote un effet de sens particulier. C'est dans ce contexte que nous avons traduit quelques phrases par Google traduction et Systran. Nous avons constaté l'impossibilité de réaliser les mêmes effets de sens produits par le choix délibéré de « à », « de » ou de l'infinitif pur. Nous avons relevé beaucoup d'obstacles sur lesquels bute la traduction automatique du langage. Pour en améliorer la qualité, nous avons en premier lieu opté pour un logiciel qui s'appelle WAMP afin de créer une base de données à laquelle nous avons attribué le nom « Traduction ». En second lieu, nous avons développé un programme en Java qui va utiliser cette base de données pour que la traduction automatique des constructions précitées soit parfaite et satisfaisante. Ce travail compte remédier aux lacunes des traducteurs automatiques qui se contentent de traduire littéralement l'énoncé.

Traduction financière et nouvelles technologies de l'information et de la communication (NTIC) : Cas de la « monétique » : moyens de paiement électronique

Leila LAIMECHE

Département de Traduction - Université Badji Mokhtar, Annaba – Algérie

lila.traduction@yahoo.fr

Résumé : Les vagues d'innovations, qui ont marqué le domaine des technologies de l'information et de la communication durant les dernières décennies, ont profondément modifié la nature des services bancaires et financiers. Ces technologies, entendues comme un large ensemble regroupant l'utilisation de bases de données de plus en plus riches, de moyens de transmission et de paiement électroniques, de centres d'appels téléphoniques, d'équipements en informatique et téléphonie mobile et surtout de réseaux de télécommunications ouverts, ont renouvelé la relation entre les banques et leurs clients ainsi que les prestations financières et les métiers bancaires.

Au fil des années, la terminologie relative aux systèmes de paiement s'est régulièrement affinée en fonction de l'évolution des infrastructures de paiement et de règlement. Les avancées technologiques soulignent l'importance de l'emploi cohérent des termes nouveaux, que nous soyons ou non des techniciens. Et comme dans la plupart des disciplines, la terminologie des paiements a aussi été enrichie par plusieurs études analytiques, sous forme de concepts et de termes, mais surtout d'abréviations. Cette terminologie souvent opaque et consistante pose beaucoup de problèmes en matière de traduction, notamment des langues étrangères vers la langue arabe. En fait, la plupart des néologismes, fruit de la monétique, n'ont pas souvent de correspondants ou d'équivalents dans la langue cible. Le terme « monétique » serait un parfait exemple de ce problème de traduction. « Monétique » vient de « monnaie » et « informatique » et renvoie aux moyens de paiement électronique susmentionnés. Le métier lié à l'activité monétique, s'appelle "monéticien" (pas d'équivalent en anglais ni en arabe). Des abréviations comme « DAB », « GAB », « ATM », « TPE », « CIB » et des termes tels que « Carte Visa », « SmartCard » « MasterCard », « PayPal » utilisés dans les langues française et anglaise viennent s'ajouter au lexique de la monétique et compliquer, par la suite, la tâche du traducteur spécialisé dans la traduction financière qui doit investir dans la langue cible pour trouver des solutions adéquates. La traduction de ces termes et abréviations monétiques se fait généralement grâce à des techniques précises comme « l'emprunt », « le calque » et « la traduction littérale », qu'on va essayer d'expliquer et d'illustrer en se basant sur un échantillon de lexiques ou glossaires (français-anglais-arabe) relatifs à la monétique.

Les difficultés et les limites de la traduction automatique de la métaphore dans le discours politique (Français / Arabe)

Marwa Ammari

Laboratoire Langage et Traitement Automatique (LLTA), Université de Sfax

ammariamrwa20@yahoo.fr

Résumé : La métaphore est une figure de style par laquelle on détourne le sens propre d'un mot. Ce sens se diffère d'une culture à une autre. Ainsi, le traducteur, en traduisant une métaphore, doit prendre en considération cette condition pour que sa traduction soit cohérente et possible. Il doit traduire dans une autre langue le contenu et les connotations de cette figure. Néanmoins, cette transposition pourrait poser des problèmes au niveau de l'équivalence surtout dans la traduction automatique de la métaphore. Les logiciels de la traduction automatique rencontrent des difficultés à connoter et décoder une phrase métaphorique. Ainsi, dans cette recherche, nous mettrons l'accent sur les difficultés de la traduction automatique de la métaphore dans le discours politique, du français vers l'arabe, en posant ces questions :

Comment peut – on traduire une métaphore ? Comment peut-on aider l'outil informatique à la détecter dans un discours politique ? Et quelles sont les limites de cette traduction automatique ?

Mots clés : métaphore- discours politique- traduction automatique - langue française - langue arabe – les logiciels de la traduction.

Mind Mapping in Classroom: Inspire and Aspire with Coggle

Bootheina Majoul

Institut Supérieur des Langues Appliquées et d'Informatique de Béja, Université de Jendouba

bootheina@yahoo.fr

Abstract: In this digital era, it has become very difficult for a teacher to compel students to read, write and think because technology invaded our teaching spaces and provided learners with ready-made ideas. Indeed, when misused technology could be a threat to learning and teaching alike. It is then part of the hard mission of the teacher to always think about alternatives to preserve learning and transmit knowledge to these digital natives.

Mind mapping proves to be one of the teaching tools that could help a teacher inspire students and enhance their creativity. This paper will highlight the efficiency of mind mapping in/outside classroom and focus mainly on the use of Goggle in different disciplines.

Tools for Technology-Enhanced Formative Assessment

Asma Gaaloul

Institut Supérieur des Etudes Appliquées en Humanités de Tunis

asma.maaoui@gmail.com

Abstract: Information technologies have been recognized as effective tools in the implementation of assessment for learning. This workshop aims at showing a number of web-tools that can facilitate the implementation of Formative Assessment. It will emphasize the specific role technology plays in enhancing learners' engagement with feedback to regulate their learning process through online formative interaction and collaboration. Thus, this workshop will hopefully raise English as a Foreign Language (EFL) teachers' awareness about the need for the integration of communication technology in assessment for learning in formal and informal settings.

Keywords: Formative Assessment, information technologies, formative feedback, collaboration, self-regulation.

L'image et son discours : analyse pragmatique d'une photo

Hajer Mzoughi
Laboratoire LLTA, Université de Sfax

Résumé : Communiquer par l'image est aussi une forme de discours. Dans ce projet de communication, nous proposerons une analyse descriptive et pragmatique d'un élément iconique qu'est l'image, accompagné d'un élément linguistique. Cet élément iconique représente l'image d'une réalité sociale opposée implicitement à une autre. Outre l'image, le texte est aussi l'expression d'un point de vue.

Notre objectif sera de démontrer, dans le cadre des théories de la communication et du langage audio-visuel, les processus de signification pilotant le travail de l'interprétation, de construction de sens et de compréhension de cette donnée iconique, considérée comme un véritable discours social. Celle-ci constitue des informations et est susceptible de transmettre un message ciblé pour l'auditeur, exprimant une certaine prise de position de son producteur. Dès lors, nous focaliserons notre analyse sur l'importance du cadre contextuel dans la construction du sens, en focalisant sur le savoir et les connaissances mises à contribution pour l'interprétation et la construction du sens du message que l'image véhicule. Nous évaluerons ensuite la pertinence de l'élément iconique en termes de transparence/opacité.

Mots clés : élément iconique- élément linguistique- signifiant/ signifié de connotation- rapport de complémentarité- inférence- pertinence.

L'Apport du traitement automatique l'analyse du discours : Mesure de l'objectivité du discours journalistique

Jamil GHOUAIDIA
Université de Carthage
ghouaidia.jamil@gmail.com

Résumé : Notre contribution s'inscrit dans le cadre de nos travaux d'analyse du discours journalistique. Elle aura pour objet d'interroger le rôle des méthodes de « fouille des textes », qu'assure le logiciel Tropes dans l'appréhension de la problématique de l'objectivité à partir du traitement automatique des articles qui relèvent de la rubrique politique du quotidien le Temps (2011). Il s'agit de montrer comment exploiter des techniques du TAL pour explorer automatiquement un corpus de presse et en traiter les spécificités linguistiques. Notre étude de l'objectivité du texte journalistique nous a poussées à prendre en compte les différentes pistes possibles pour la concevoir. Nous chercherons à articuler, à partir de cet exposé, traitement automatique et analyse du discours, méthodes statistiques et méthodes symboliques afin de parvenir à expliciter un phénomène langagier un peu complexe à l'instar de l'objectivité. Notre contribution s'insère dans la fouille des données textuelle et s'articule autour de deux axes. Nous exposerons, dans un premier temps, notre méthodologie, notre logiciel et décrirons notre corpus. Puis, dans un second temps, nous soulignerons l'importance des données fournies dans le calcul de l'objectivité du texte de presse et ce par la conjugaison des méthodes statistiques et de la description linguistique.

Le discours électronique extrémiste

Yosra Sayah

Faculté des Lettres et des Sciences Humaines de Sfax et Membre de LLTA

Yosra_gbf91@hotmail.fr

Résumé : L'émergence de l'internet dans la vie quotidienne présente, pour de nombreux chercheurs s'inscrivant dans le domaine des sciences du langage, une opportunité de se lancer dans des travaux sur de nouveaux corpus extraits de ce réseau. Sites web, courriels, youtube, Facebook, Twitter, ... constituent, à nos jours, un champ de recherche fertile mais épineux. La présente contribution s'intéresse aux discours électroniques et en particulier au discours extrémiste sur les réseaux sociaux ainsi qu'au rôle que peut jouer l'analyse automatique du discours dans l'étude de ce type de corpus.

Mots clés : Discours électronique, discours extrémiste, réseaux sociaux, corpus, analyse automatique du discours.

La polysémie du marqueur évidentiel «évidemment» dans le discours de Chirac

Adel Mnefgui
Laboratoire LLTA, Université de Sfax
mnefguiadel2@yahoo.fr

Résumé : Une des plus importantes fonctions du langage est de transmettre de l'information qui peut être parvenue des sources variées. Le locuteur l'a obtenue par observation, par inférence, à partir d'indices, il peut l'avoir recueillie comme une rumeur, ou encore l'avoir reçue en rêve.

Au moment de produire un énoncé le locuteur marque alors dans son énoncé la source du savoir que cet énoncé est censé communiquer. En effet la transmission de l'information est l'objet primordial de chaque acte de parole. Chaque langue dispose certes des outils linguistiques et lexicaux pour préciser la source de l'information. Dans ce contexte, on va focaliser notre étude sur le connecteur évidentiel « évidemment » comme une expression langagière qui apparaît dans l'énoncé et qui indique la manière dont elle est obtenue une information énoncée par un locuteur et si l'information transmise dans cet énoncé a été empruntée par le locuteur à autrui ou si elle a été créée par le locuteur lui-même.

Remarquons par ailleurs que « évidemment » n'est pas un marqueur évidentiel monosémique, mais il a de multiples significations qui sont le résultat de modulations contextuelles variées. Donc il a une valeur polysémique qui donne lieu à la possibilité d'étendre et de varier le sens de l'énoncé en question

L'inférence et les modes de raisonnement : entre inférence valide et inférence valable

Mohamed Bouattour

Faculté des Lettres et Sciences Humaines, Laboratoire du langage et traitement automatique.

L.L.T.A, Université de Sfax

mohamedbouattour@yahoo.fr

Résumé : Dans le discours ordinaire, raisonner consiste, selon Aristote, à *inférer à partir de certaines choses d'autres choses, qui pourront être aussi bien des situations, des valeurs que des jugements*. Tout raisonnement, rappelle G. Vignaux¹, peut être ainsi traduit en schéma d'inférence comme le font les logiciens : «Tout A est B et tout C est A donc tout C est B.» Mais le problème, selon Vignaux, «c'est qu'on ne sait jamais ce qui vient garantir la validité d'une telle inférence. En logique, cela se fait au moyen de lois, dérivées d'un petit nombre d'axiomes, lesquels garantissent que telle ou telle inférence est légitime dans un système. Il n'en va pas de même dans le raisonnement naturel ». Dans ce dernier, il est extrêmement rare qu'un locuteur normal, en situation de communication naturelle, procède par définition, ensuite par axiome pour en inférer un jugement ou une opinion. Même dans les cas où les interlocuteurs développeraient des syllogismes, il ne serait pas toujours acquis qu'ils aient respecté les relations entre les termes (le premier et le second) dans la majeure et la mineure, ni leur impact sur la conclusion. C'est d'ailleurs la raison pour laquelle on suppose que le mode de raisonnement valide relève de la logique dans laquelle les relations d'implication sont les plus fréquentes et les plus pertinentes. En revanche, les raisonnements valables, admis ou acceptables relèvent de la rhétorique du discours quotidien où les relations d'inférence sont certes les plus utilisées. Ceci étant dit, il n'est pas sans intérêt de rappeler que même dans le discours quotidien, le locuteur a parfois recours à des règles d'implication.

¹ Vignaux, G., *L'argumentation*, Paris, Hatier, 1999.

Dictionnaire électronique de la prédication épistémique : le nom

Bouzid Abdessattar

Université de Sousse

abdessattar.bouzid08@hotmail.fr

Résumé : L'élaboration d'un dictionnaire électronique dédié au traitement automatique des langues en général, et des prédicats nominaux épistémiques en particulier, constitue un contexte applicatif qui permet de valider les différentes options théoriques d'une telle modélisation linguistique. Ce modèle électronique a comme principal objectif l'association de toutes les entrées lexicales à des informations métalinguistiques standardisées pour donner lieu à des procédures informatisées. Dans cette perspective, la langue étudiée doit être représentée d'une manière formalisée. Autrement dit, la structuration des bases de données d'un dictionnaire électronique exige que les faits linguistiques soient représentés d'une façon explicite et normalisée.

Pour atteindre cet objectif, un fait de langue est étudié selon deux dimensions possibles à savoir une première dimension théorique et une deuxième dimension applicative. Cette interaction entre ces deux dimensions permet de préciser l'importance de la théorie des classes d'objets, théorie de Gaston Gross.

Mots clés : Prédicat nominal, dictionnaire électronique, classes d'objets, entrée lexicale, emploi prédicatif, nom épistémique.

L'inférence automatique (le moteur de recherche : Google) et l'inférence humaine

Berfifa Abir

Faculté des lettres et des sciences humaines de Sfax,

Laboratoire de traitement automatique (LLTA), Université de Sfax

berfifaabir@yahoo.fr

Résumé : C'est vrai que l'informatique et l'internet facilitent et modernisent les recherches linguistiques. Elles participent à la naissance des créations linguistico-informatiques honorables et enrichissantes, mais cela n'empêche pas d'affirmer que l'intelligence artificielle ne peut traiter plus ou moins que le côté codique. Autrement dit, l'ordinateur est incapable de lire, réfléchir sur les faits linguistiques et déduire le sens. Par contre, l'être humain en est apte parce qu'il possède des connaissances et une faculté qui lui permet de raisonner et de prendre des décisions. C'est dans ce cadre que s'inscrit la problématique de notre présent article qui porte sur l'inférence. Les moteurs de recherche ne peuvent fonctionner que par mots clés. Ils sont incapables de distinguer les divers faits linguistiques et de calculer le sens de la phrase contrairement à l'Homme qui infère facilement selon les différents contextes. Nous avons mis en lumière ce problème à partir de la comparaison entre l'inférence chez les interlocuteurs humains, d'une part, et l'internaute face à Google, d'autre part, grâce à un corpus des questions/ réponses, bien précis.

Mots clés : Inférence, Homme, question, machine, sémantique, réponse, moteur de recherche (Google) et pragmatique.

Défigements à *La Une* de la presse écrite tunisienne

Boughammoura Amine
Faculté des lettres et des sciences humaines de Sousse
Laboratoire Langage et traitement automatique (LLTA)
Amine-b1@live.com

Résumé : La spécificité des titres de presse réside dans leurs tâches pragmatico-communicatives. Dans cette étude, nous examinons les textes journalistiques numériques tunisiens en focalisant notre attention sur les titres de « unes » jouant sur les mots.

Les questions auxquelles nous tenterons de répondre sont les suivantes. Par quels procédés et avec quels effets de sens les jeux de mots des titres de « unes » sont-ils réalisés ? Sur quelle base linguistico-culturelle prennent-ils naissance ?

Dans un premier temps nous définissons la notion de défigement. Dans un second temps, nous nous intéressons au repérage et à l'analyse de ces jeux de mots. Enfin, nous montrons que le traitement automatique des ces occurrences défigées est problématique.

Mots clés : presse numérique, défigement, procédés de défigement, base linguistico-culturelle, traitement automatique

La prédication du SN déterminé par l'adjectif démonstratif dans le cadre anaphorique

Sana Abdelhamid

Laboratoire Langage et traitement automatique (LLTA)

sanaabdelhamid@yahoo.fr

Résumé : Notre travail s'insère dans le cadre général de l'étude du phénomène anaphorique en vue d'améliorer la traduction automatique des anaphores. Dans cette communication, nous nous proposons de traiter un type particulier d'anaphore à savoir l'anaphore démonstrative, à travers un corpus que nous avons puisé dans des textes littéraires et des articles de presse et ce dans le but de faciliter le traitement des expressions anaphoriques à déterminant démonstratif. Nous nous proposons de démontrer que le SN déterminé par l'adjectif démonstratif, au-delà de sa valeur de reprise (étant donné que l'anaphore est conçue comme une relation essentiellement d'ordre textuel qui mettrait en jeu un antécédent et un anaphorique) est un moyen de prédication, notamment dans le cas des anaphores attributives, qui reprend au moyen d'un prédicat un élément qui soit ou non textuellement présent (le cas des anaphores présentant une ellipse de l'antécédent). Cette prédication peut être de différents ordres : identificationnelle, évaluative, prédicationnelle ou encore implicite dans le cas des « prédications présentées comme connues que l'on trouve chez Milner (1982) sous la notion d'« anaphore présuppositionnelle ». Elle renvoie à un contenu discursif d'un ensemble de propositions qui viennent d'être énoncées tout en établissant une relation de type « être » entre les N des SN démonstratifs et leur source. Il s'agit d'un enrichissement informatif qui répond au principe de **continuation** dans la chaîne discursive des procédés prédictifs.

Emotions et interactivité sur Twitter

Najeh ELOUNI

Laboratoire Langage et Traitement Automatique (LLTA) / Université de Sfax
Centre Interlangues, Texte, Image, Langage (TIL) / Université de Bourgogne, Franche-Comté
najhelouni@gmail.com

Résumé : Le présent travail met l'accent sur deux aspects fondamentaux sur Twitter à savoir l'interactivité et la dimension émotionnelle. Ces deux dimensions sont traduites par divers procédés verbaux et non-verbaux. L'écriture sur Twitter est très contraignante aussi bien au niveau de la longueur des messages limités à 140 caractères qu'au niveau de leur fonctionnement conditionné par quatre opérateurs organisant la communication et l'interaction (hashtag#, @, http://, RT). Ces opérateurs orientent le lecteur à l'extérieur du tweet régulant ainsi les interactions à distance que créent les tweets. L'opérateur@ constitue le parfait exemple véhiculant cet aspect interactif. Une double interactivité est ainsi installée à la fois entre les utilisateurs établissant un réseau d'interactions et de relations interhumaines, mais aussi entre Twitter et les autres réseaux sociaux.

Mots-clés : tweets, émotions, interactivité, hypertextualité, opérateur @.

Description et méthodologie de traitement d'un corpus de tweets : cas de pragmatèmes d'affect

Najwa Gharbi

Faculté des Lettres et des Sciences Humaines, Université Grenoble Alpes

najwa99.gharbi@gmail.com

Résumé : Notre objectif est d'étudier la notion de pragmatèmes d'affect dans des environnements médiatisés par les réseaux, à partir d'un corpus déjà constitué dans le cadre de projet CoMeRe (Communication médiée par les Réseaux). Ce projet a élaboré une banque de corpus de nouveaux moyens de communication tels que les tweets. Il importe donc de présenter ce projet en explicitant la méthode de construction des corpus. Cet article se propose tout d'abord, d'examiner l'utilisation des pragmatèmes dans un type particulier du corpus de tweets. Ensuite, nous allons préciser notre méthodologie de traitement et d'extraction de données à travers le logiciel AntConc.

Etude des collocations verbales de sentiment en français et en arabe : à visées applicatives en traitement automatisé du lexique de sentiment

Safa Zouaidi

LLTA, Université de Sfax

safa.zouaidi@doctorant.univ-grenoble.fr

Résumé : Mobiliser des corpus numériques pour étudier des phénomènes linguistiques est indispensable (Blumenthal et Hausmann, 2006). Notre étude s'insère dans le cadre de la linguistique des corpus (Sinclair, 1991). La question principale concerne les correspondances entre le français et l'arabe dans l'expression de sentiment appartenant aux champs sémantiques de la surprise, la colère, l'admiration et la jalousie.

Pour les verbes de sentiment, certains auteurs en proposent un classement en s'appuyant sur des critères syntaxiques (Harris, 1988 ; Gross, 1995, etc.). D'autres proposent une répartition des verbes en se fondant sur des traits sémantiques homogènes (Mathieu, 2000, etc.). Nous proposons ici une troisième voie qui consiste à relier systématiquement la sémantique (les dimensions sémantiques véhiculées par les collocatifs) à la syntaxe (les constructions grammaticales récurrentes).

Pour cela, nous solliciterons deux corpus journalistiques numériques comparables et équilibrés : EmoBase (projet Emolex 120 M de mots) et ArabiCorpus (130 M de mots). En effet, l'interface EmoBase est interrogeable en ligne à partir de certaines méthodes d'interrogation (Kraif, 2002) et permet d'obtenir des résultats détaillés concernant toutes les occurrences d'un pivot donné dans le corpus. Quant à l'interface ArabiCorpus, elle est, à première vue, moins ergonomique que celle d'Emolex. Elle ne dispose pas de structure d'annotation ni, d'ailleurs, d'indexation préstructurée, ce qui peut rendre l'opération de repérages des collocations en question beaucoup plus chronophage.

Ces corpus nous permettront de nous interroger, en particulier, sur les correspondances syntaxico-sémantiques dans les combinaisons verbales entre les deux langues. Nous chercherons plus précisément à savoir si des verbes appartenant aux mêmes types de sentiment dans les deux langues attireraient les mêmes accompagnateurs ou non. Nous choisirons donc de fonder notre étude sur l'analyse de la combinatoire syntaxique et lexicale (Novakova et Tutin, 2009). Nous nous baserons aussi sur la théorie du *Lexical priming* (Hoey, 2005), qui nous permettra d'examiner les relations syntaxiques privilégiées des combinaisons verbales de sentiment et de vérifier si elles sont liées aux dimensions sémantiques qu'elles peuvent éventuellement véhiculer.

الاسم والفعل في التعبير عن الألم: دراسة من خلال آلية المسح العرفني

الحبيب المقدّيني

المعهد العالي للّغات التّطبيقيّة والإعلانيّة بباجة، جامعة جندوبة

elmkadmini@gmail.com

ملخص العمل: الألم ظاهرة كويّة لازمة للوجود الإنساني، ولها تداد في حيط الفرد الاجتماعي والثقافي، وهي كذلك تجربة وثيقة الاتّباط بالّلغة لألم تجربة استبطنة ذاتيّة لا فهمها إلّا من خلال التلقظ بها. فالّلغة ويلة أليّة عتمدها لإظهار هذه التّجربة والإفصاح عنها. تتجلى هذه الأهميّة في التعريفات العالميّة للألم، وأن ذلك التعريف الذي وضعته (الجمعيّة العالميّة للألم² IASP). وصّه:

"الألم تجربة زعجة ذات طبيعة سية أو عنويّة، يعود ببها إلى تلف واقعي أو ممكن في أحد الأسجة من الجسم. ويمكن أن يوصّف الألم من خلال ما يحيل على التّلف التّسجيّ ويُتعلّق الأمر بألم فسيّ".

فالّلغة من هذا المنظور ذات قيمة ركزيّة في تحديد الألم وتقييمه بالإضافة إلى دورها الأليّ من حيث كيات ويلة التّواصل الطّبيعيّة بين الأفراد. وقد أصبحت "لغة الألم"³ جالا تطبيقيا للّغات اللّسانيّة بمختلف توجّهاها وظريّاتها. وقد التّ المقاربة اللّسانيّة العرفنيّة⁴ ضها من هذا المجال في سائل مختلفة حويّة ودلاليّة نها الاتّعلاقة في التّعبير عن الألم ونه ظريّة الأطر ونها الأليات النّحويّة العرفنيّة في التّعبير عن هذه التّجربة. وسعى في هذا العمل إلى دقة التّعبير العربيّ عن الألم من خلال إدنى الأليات العرفنيّة المنعكسة في النّحو والمقولات النّحويّة⁵ المعتمدة في التّعبير عن هذه التّجربة؛ وعني هنا آليّة المسح⁶ ودورها في تشكيل ألمات الإبناء⁷ المختلفة (الاميّة والفعلية)، وما تعكسه من أس فهويّة في تصوّرنا لهذه التّجربة وما يحيط بها، وفي تعلّقنا بها.

المفردات الأساسيّة: لغة الألم، حو عرفنيّ، جال فهويّ، المسح، المسح التّتابعيّ، المسح المجمل، المقولات النّحويّة، العمليّة الرّتابيّة، العمليّة اللّازماتيّة

² (IASP) International Association for the Study of Pain

³ pain language

⁴ cognitive linguistics

⁵ grammatical categories

⁶ scanning

⁷ construal (s)

اللسانيات الحاسوبية : من الوصف إلى التوصيف

اللغة والتكنولوجيا في العلاقات العامة : آلية الإقناع وبناء السمعة الرقمية

إلي مود، باثثة في المعهد العالي للدكتوراه في العلوم الإنسانية والاجتماعية والإعلام- الجامعة اللبنانية

sallyhammoud01@gmail.com

غسل مراد، أستاذ اللسانيات الحاسوبية - الجامعة اللبنانية

ghassa.mourad@ul.edu.lb

مقدمة: تعتبر العلاقات العامة لغة الوصل والتواصل بين المؤسسات والمجتمع الواسع، وبين المجتمعات الأخرى. وتتداخل في تركيبها وظائفها عدة تشمل العلاقات مع وسائل الإعلام والاتصال والبحث والإنتاج والتقييم والإدارة والتخطيط والقضايا العامة وغيرها. العلاقات العامة توجه وسائل اتصالية إلى الجماهير من المستهلك، كونه، مجتمع يابى أو اقتصادي وقادة الرأي العام. العلاقات العامة ليست فقط جزء من السياقات التسويقية والترويجية، بل تشمل أيضاً تراتيجيات جماهيرية (Mass) تهدف إلى تشكيل لوكيات فردية وجماعية قادرة على خلق أي عام من خلال قدرتها على الإقناع.

فالربط بين دور العلاقات العامة وعملية الإقناع في تغيير السلوكيات البشرية شكلت جسر العنق بين مفهوم البروباغندا للإطلاق نحو علم العلاقات العامة وتأثيرها في تغيير الثقافات داخل المجتمعات. ومن مجالات عمل الإقناع في العلاقات العامة هناك الأساليب الثلاثة الآتية التي تركز عليها عملية التسويق:

- تسويق منتجات
- تسويق أفكار وقيم
- الوعي للأفراد (أو ما يعرف بالتوعية أو الإعلام الملتزم)

تساؤلات حول دور علامات الترقيم في الترجمة البشرية والآلية

غسل □□ راد

□ تاذ دكتور □ في اللسانيات الحاسوبية والاعلام الرقمي، □ مركز علوم اللغة والتواصل وكلية الاعلام - جامعة الليبية

ghassan.mourad@ul.edu.lb

ملخص : تطرح علامات الترقيم والتنقيط عدة تساؤلات لنوعية فهمها ودورها بالخاص عندنا يتعلق الأمر بالترجمة بشرية كانت أم آلية. وإن هنا يأتي السؤال الأول هل فعلاً □□ الممكن □□ سمي ذلك ترجمة كما هو الحال بالنسبة للمفردات... أم □□ هذه العملية ليست إلا قل للحالة الوظيفية لعلامات الترقيم □□ لغة □□ صد □□ الى لغة هدف. □□ طلاقاً □□ ذلك لهذا البحث عدة ابعاد لخصها بالآتي:

اولاً، هل □□ الممكن □□ ترجم علامات الترقيم والتنقيط واذا □□ جرى □□ هي المعايير التي □□ المفترض اعتمادها بالخاص □□ وظيفة هذه العلامات تختلف □□ لغة الى اخرى. □□ انها □□ هي □□ عمدة □□ عيانية في اللغات الاوروبية □□ ثلاً، □□ انها □□ هي غير □□ عمدة ولا تتبع اي □□ عيانية دلالي او □□ حوي كما في اللغة العربية، إذ يبقى ذلك على □□ الكاتب. لذلك □□ نعالج في هذا الموضوع هذه الاشكالية لنرى □□ اذا كانت هذه المسألة تؤدي الى اللبس اللغوي والى عدم اقبال المعنى بشكل صحيح.

ثانياً، تشكل □□ مسألة ترجمة علامات الترقيم والتنقيط عائق □□ ام تغذية الذاكرة الترجمة المعتمدة في الترجمة الآلية المساعدة للمترجم. ثالثاً، اذا □□ المترجم يعتمد على الروابط لكي يستخدمها □□ ك □□ العلامات المترجمة □□ اللغات الاخرى، "□□ اذا عن المترجم الآلي كمترجم غوغول □□ ثلاً الذي يعتمد على ترجمة الجمل.

□□ ابعاً تطرح □□ مسألة ترجمة علامات الترقيم اشكالية تعريف الجملة □□ الناقية الدلالية. فإذا كانت النقطة تحدد □□ هاية الجملة، و □□ الجملة لها □□ عن □□ حدد كما هو □□ د في كتب القواعد، إذا □□ العمل عند □□ ترجم جملة باثنتان □□ او اثنتان □□ بوا □□ دة؟ هل هذا يطرح اشكالية تعريف الجملة كمكو □□ لغوي □□ عنوي داخل النص؟

□□ ينطبق على الجملة □□ الممكن □□ ينطبق على □□ هو اقل □□ الجملة كاشباه الجمل □□ ثلاً، □□ ينطبق على الفقرة كمكو □□ دلالي يحتوي على عدة جمل.

لمحاولة الإجابة عن هذه التساؤلات □□ نعتد على □□ قافية □□ دوية □□ ترجمة □□ الفرسية الى العربية.

أخيراً، لا بد □□ طرح □□ مسألة علامات الترقيم والتنقيط باللغة العربية وكيفية □□ تحدا □□ ها، ك □□ لها لم تشكل □□ نى □□ هاجساً لغوياً يهم الالسنين. ولكن ك □□ علم فعلاً □□ علامات الترقيم والتنقيط □□ في بناء المعنى، ولم يعد ب □□ تطاعتنا عدم اخذه بالاعتبار، □□ كل □□ لجهة النص المصدر، او لجهة النص الهدف.

Invités du colloque Infol@ngues 2018



Prof. Musapha Jarrar
Université de Birzeit
Palestine

Associate professor of Computer Science, at Birzeit University in Palestine. He was a Fulbright visiting professor at the University at Buffalo in the USA (2016-2017), a Marie Curie fellow at the University of Cyprus (2007-2009), and a senior research scientist at Vrije Universiteit Brussel (1999-2007), where he completed his Masters and PhD. Prof. Jarrar published +75 articles and refereed reports in the areas of Ontology Engineering, Lexical Semantics, Natural Language Processing, Semantic Web, and Databases. He chaired 23 international workshops, a PC member of +100s journals/conferences, a coordinator/manager of 25 large EU project, a full member of the IFIP2.6 on Database Semantics, the IFIP2.12 on Web Semantics, and the UN ESCWA Technology Centre Board of Governors, among others. Prof. Jarrar is also the founder of both Sina Institute For Knowledge Engineering And Language Technologies, and the Palestinian eGovernment Academy. He also was voluntarily serving as an advisor at the Palestinian ministry of Telecom & IT for e-government topics, where he developed and chaired the Palestinian e-Government Interoperability Framework (called Zinnar).



Prof. Lahcen Amor
Université
de Badgi Mokhtar
Algérie

الأستاذ الدكتور عمر لحسن هو استاذ التعليم العالي متحصل على شهادة دكتوراه دولة في اللسانيات بعنوان "الخصائص المعجمية والدلالية في ديوان جرير". بدأ التدريس بالجامعة منذ 1982 اي لديه 36 سنة خبرة.

درّس المواد اللسانية بقسم اللغة العربية وادابها بالإضافة الى الإشراف على عشرات بحوث التخرج (اجازة، ماجستير و دكتوراه) في اللسانيات واللغة العربية.

تقمص العديد من المسؤوليات الإدارية والتربوية، من اهمها رئيس قسم اللغة الفرنسية، كلف بمهمة التوثيق لدى رئيس جامعة عنابة، نائب مدير معهد اللغة العربية وادابها، ...

حاليا هو مدير مخبر اللسانيات و اللغة العربية، شارك في العديد من من الملتقيات العلمية الوطنية و الدولية. كما انه نشر العديد من المقالات وترجم العديد من المؤلفات.



Dr. Faiza Derbel
Université de La
Manouba
Tunisie

Faiza Derbel is assistant Professor of English and Applied Linguistics at the English Department, Faculty of Letters, Arts and Humanities of the University of Manouba. She teaches graduate-level courses and conducts and supervises research in applied linguistics with special interest in instructed second language acquisition (ISLA), teacher education, and language learning/teaching with technology.