


25
Juin

DOCTORIALES LILPA

Journée des doctorantes et doctorants de l'unité de recherche, conviviale et bienveillante.
Présentation de travaux, discussions, échanges...

Amphi Beretz
(et en ligne)

 Laboratoire

Linguistique, Langues et Parole | LILPA | EA 1339

Université de Strasbourg

Programme de la journée

9h45 - 10h00 Accueil des participantes et participants

10h00 - 10h15 Pr. Rudolph SOCK : Mot de bienvenue



10h15

Anaïs KOEHLER

Propriétés des code-switchings (CS) des locuteurs dialectophones de l'Alsace contemporaine

10h45

10h45

Ioana BUHNILA

Annotation automatique des termes médicaux pour l'identification des marqueurs de reformulation médicale

11h15

11h30

Gercélia BATISTA

Le dialogue entre l'objet et sa réalité environnante, une interaction essentielle à l'activité définitoire en terminologie juridique

11h50

11h50

Cindy COELHO

Il était un petit Paul : bégaiement et plurilinguisme au Luxembourg

12h10



Pause déjeuner

14h00

Chang GUO

14h30

Étude contrastive des chaînes de référence en français et en chinois — Étude de cas à partir d'un corpus parallèle composé de Jean-Christophe de Romain Rolland et de sa traduction par Lei Fu

14h30

Oti AYENBI

15h00

Minority Language and Language Endangerment: A case study of Itsekiri language in Nigeria

15h15

Erin McINERNEY

15h45

L'éthique dans la recherche sur les réseaux sociaux : le consentement éclairé sur Instagram

15h45

Amandine ROBERT

16h15

Voilà en français parlé contemporain : du déictique au marqueur de discours

16h15 - 16h35 Pr. Jean-Christophe PELLAT : Mot de synthèse

16h45 - 17h45 Discussion conviviale doctorantes, doctorants et masters



Inscription obligatoire via ce lien. Plus d'informations sur le site LiLa.



***Minority Language and Language Endangerment: A case study of Itsekiri language in
Nigeria***

Oti AYENBI – 5^e année de doctorat

The study investigates the question of language endangerment in a minority language setting. It analyses language shift in consonance with linguistic practices and behaviour of speakers of Itsekiri language in Nigeria. The objectives are to ascertain the causes of language shift in the Itsekiri community, to uncover and understand reasons behind the phenomenon of non-inter-generational transmission of Itsekiri language and the seeming preference for English language. To attain these objectives, the study seeks to address the following questions: What are the causes of language shift in Itsekiri community? What are the factors responsible for non-intergenerational transmission? What is the role of English and Pidgin in the Itsekiri community? What could be the possible strategies to reverse language shift and maintain Itsekiri language? There are three hypotheses to this study. Firstly, Itsekiri is endangered and risks extinction due to the phenomenon of language shift and non-transmission from one generation to another. Secondly, proficiency in English is prerequisite to success in academics and an effective means of social mobility. Thirdly, linguistic adaptation underpins language shift. Based on theoretical frameworks underlying Language Endangerment proposed by Fishman J., (1991), Austin & Sallabank (2011) and Whaley & Grenoble, the study uses qualitative methodology with ethnographic approach to find answers to a myriad of underlying questions surrounding implied assumptions, with a view to ascertaining what is going on linguistically in the community.

***Le dialogue entre l'objet et sa réalité environnante, une interaction essentielle à
l'activité définitoire en terminologie juridique***

Gercélia BATISTA – 2^e année de doctorat

Mots-clés : Langues spécialisées - Définition terminologique – Terminologie juridique – Terminographie

Partant d'une analyse fondamentale de la définition dans la langue générale et dans les langues spécialisées, selon les principes de la définition terminologique énoncés par Pearson, Cabré et Van Campenhoudt, entre autres, cette communication a pour but de proposer une réflexion sur la contribution que les données de la réalité entourant l'objet devant être décrit peuvent offrir à l'activité définitoire en général et dans un domaine spécialisé, celui de la réglementation de la fin de vie, en particulier.

Nous nous appuyerons sur les constatations établies dans le cadre d'un travail réalisé sur un *corpus* d'un peu plus d'un million de mots, constitué de textes de lois, de jurisprudence et de doctrine en allemand, français et portugais du Brésil, pour essayer de démontrer les incidences positives que les ajouts issus de ces trois réalités du droit peuvent avoir sur l'activité définitoire du terminologue.

Nous nous servirons d'un exemple extrait de notre base terminologique, non seulement pour attester de l'importance de la relation entre la conception du *definiendum* comme notion et unité terminologique et les composantes de sa réalité *de facto* (le fait, l'acte, le sujet) et *de jure* (la loi, la jurisprudence, la doctrine) pour l'élaboration d'un *definiens* approprié.

Annotation automatique des termes médicaux pour l'identification des marqueurs de reformulation médicale

Ioana BUHNILA, 3^e année de doctorat

Mots-clés : annotation automatique, termes médicaux monolexicaux et polylexicaux, SIFR-BioPortal, marqueurs de reformulation, corpus de reformulation

Les textes de spécialité médicale posent des problèmes de compréhension au grand public à cause de la complexité des termes médicaux. Vue l'importance croissante de rendre les recherches médicales compréhensibles par le grand public, nous nous concentrons dans ce travail sur l'identification automatique des termes médicaux et leurs reformulations. Nous exploitons le corpus français comparable écrit du domaine médical de littérature scientifique et de textes de vulgarisation CLEAR (Grabar et Cardon 2018). Nous utilisons l'annotateur SIFR-BioPortal (Tchechmedjiev et al. 2018) avec la base terminologique médicale SNOMED-3.5VF (Côté et al., 1998) pour identifier automatiquement les termes médicaux simples et polylexicaux. Nous recherchons dans leur contexte des marqueurs de reformulations avec les patrons lexicaux, de type « signifie », « veut dire » « autrement dit » (Grabar et Eshkol-Taravella 2016a), « l'équivalent de », « ou encore » (Antoine et Grabar 2016). Nous rajoutons des marqueurs observés dans le corpus « est une maladie », « est une affection » et les variants du verbe « désigner », « défini/e » et « définition ».

Nous présentons les résultats de l'analyse de corpus réalisée avec l'outil de textométrie TXM (Heiden et al. 2010). Nous avons trouvé 2 220 phrases qui contiennent des reformulations des termes médicaux parmi 4 662 phrases annotées et validées. L'analyse des données nous a permis d'identifier des nouveaux marqueurs de reformulation, de type « y compris », « comme », « notamment », « par exemple » et les parenthèses. Nous annotons les

reformulations identifiées selon la classification de fonctions pragmatiques de Grabar et Eshkol-Taravella (2016) (définition, explication, paraphrase etc.)

Références bibliographiques :

ANTOINE, E. & GRABAR, N. (2016). Exploitation de reformulations pour l'acquisition d'un vocabulaire expert/non expert. In TALN 2016: Traitement Automatique des Langues Naturelles. Paris, France. <https://hal.archives-ouvertes.fr/hal-01426816>.

CHAROLLES, M. & COLTIER, D. (1986). Le contrôle de la compréhension dans une activité rédactionnelle : l'exemple des reformulations paraphrastiques. *Pratiques* 49 (1) : 51-66. <https://doi.org/10.3406/prati.1986.2450>.

COTE, R. (1996). Répertoire d'anatomopathologie de la SNOMED internationale, v3.4. Université de Sherbrooke, Sherbrooke, Québec.

ESHKOL-TARAVELLA, I., BAUDE O., MAUREL D., HRIBA L., DUGUA C., & TELLIER I. (2011). Un grand corpus oral « disponible » : le corpus d'Orléans 1 1968-2012. *Traitement Automatique des Langues* 53 (2): 17-46.

GRABAR, N., & ESHKOL-TARAVELLA, I. (2016). Disambiguation of occurrences of reformulation markers c'est-à-dire, disons, ça veut dire. *JADT 2016 : 13ème Journées internationales d'Analyse statistique des Données Textuelles*, 1-13. Nice, France.

GRABAR, N., & HAMON T. (2015). Extraction automatique de paraphrases grand public pour les termes médicaux. *22ème Traitement Automatique des Langues Naturelles*, 14. Caen, France.

GRABAR, N., & HAMON, T. (2016). Exploitation de la morphologie pour l'extraction automatique de paraphrases grand public des termes médicaux. *Traitement Automatique des Langues, Varia*, 57 (1) : 85-109.

GRABAR, N. & CARDON, C. (2018). CLEAR - Simple Corpus for Medical French. In *Proceedings of the 1st Workshop on Automatic Text Adaptation (ATA)*, 3-9. Tilburg, the Netherlands :

- Association for Computational Linguistics. <https://doi.org/10.18653/v1/W18-7002>.
- HEIDEN, S., MAGUE, J-P., & PINCEMIN, B. (2010). TXM : Une plateforme logicielle open-source pour la textométrie – conception et développement. I. C. Sergio Bolasco (Ed.), Proc. of 10th International Conference on the Statistical Analysis of Textual Data - JADT 2010 (Vol. 2, p. 1021-1032). Edizioni Universitarie di Lettere Economia Diritto, Roma, Italy. Online.
- MAGRI, V. (2018). Marqueurs de reformulation : exploration outillée et contrastive dans deux corpus narratifs. *Langages* N° 212 (4): 35-50.
- PÉRY-WOODLEY M. & REBEYROLLE J. (1998). Domain and genre in sublanguage text: definitional microtexts in three corpora, *LREC*, p. 987-992.
- STEUCKARDT, A. (2018). Les marqueurs de reformulation formés sur dire : exploration outillée. *Langages* N° 212 (4): 17-34.
- TCHECHMEDJIEV, A., ABDAOUI, A., EMONET, V., ZEVIIO, S. & JONQUET, C. (2018). SIFR annotator: ontology-based semantic annotation of French biomedical text and clinical notes. *BMC bioinformatics*, 19(1), 405.
- VASSILIADOU, H. (2013). C'est-à-dire (que) : embrayeur d'énonciation. *Semen. Revue de sémio-linguistique des textes et discours*, no 36 (octobre). 1-14. <http://journals.openedition.org/semen/9684>. <https://doi.org/10.4000/semen.9684>

Il était un petit Paul : bégaiement et plurilinguisme au Luxembourg

Cindy COELHO, 2^e année de doctorat

Mots-clés : bégaiement, plurilinguisme, Luxembourg

Ma proposition est une présentation de mon sujet de thèse qui se base sur le modèle du concours *Ma Thèse en 180 secondes*. Ce type de concours a pour but de présenter ses travaux de recherche en simplifiant les termes de façon à ce que l'explication soit accessible et compréhensible par tout public. L'autre défi à relever est de tenir cette présentation attractive en tout juste 180 secondes et pas une de plus. C'est toutefois dans cet aspect que cette contribution s'éloigne du modèle du concours, car il ne sera pas question de se fixer une limite de temps de présentation. L'objectif visé est de présenter mes travaux de recherche de manière claire et attractive afin qu'ils soient accessibles au plus grand nombre de personnes.

Pour cette présentation, il s'agira d'une simplification scientifique de mon travail de recherche en racontant l'histoire de Paul, résident du Luxembourg qui est une personne plurilingue qui bégaié.

Paul a un problème de communication, mais cela ne se voit pas, cela s'entend, et c'est précisément ce que nous n'entendons pas qui pose problème car son discours n'est pas fluent et présente des irrégularités dans le rythme de la parole. Paul souffre de bégaiement. Mais qu'est-ce que le bégaiement ? Comment le définir ? Comment le comprendre ? Paul étant plurilingue, va-t-il bégayer de la même manière dans toutes les langues ? Les réponses à ces questions seront dévoilées dans l'histoire de Paul.

Étude contrastive des chaînes de référence en français et en chinois ---- Étude de cas à partir d'un corpus parallèle composé de Jean-Christophe de Romain Rolland et de sa traduction par Lei Fu

Chang GUO, 5^e année de doctorat

Les chaînes de référence sont définies par de nombreux auteurs comme les suites des expressions coréférentielles dans le même contexte (notamment Chastain, 1975 ; Charolles, 1988 ; Schnedecker, 1997). Cette notion est essentielle à la fois pour constituer le lien entre le monde linguistique et le monde extra-linguistique, et également pour rassurer la cohérence trans-phrastique dans le discours qui ne contient pas (beaucoup) de connecteurs logiques traditionnels. Notre présentation porte sur la question des variations qui se trouveraient entre les chaînes de référence en français et celles en chinois, en s'appuyant sur une étude de cas dans un corpus parallèle. Nous nous posons des questions dans deux aspects : quels cadres théoriques serviraient à analyser les deux langues et quelles méthodes de traitement fonctionneraient dans un corpus franco-chinois ? Par exemple, les variations inter-linguistiques s'expliqueraient par des théories comme la Théorie de l'Accessibilité (Ariel, 1990), la Hiérarchie du Donné (Gundel *et al.*, 1993) et la Théorie du Centrage (Grosz *et al.*, 1995). À l'aide d'outils informatiques, nous avons traité un corpus d'environ 52 000 mots en français et sa traduction en chinois d'environ 82 000 mots, dans lequel nous avons repéré environ 500 chaînes de référence respectivement pour chaque langue. Nous présenterons certaines variations en tenant compte de plusieurs aspects comme le type sémantique du référent, la fonction syntaxique et la forme morphologique dans les chaînes de référence des deux langues.

*Propriétés des code-switchings (CS) des locuteurs dialectophones de l'Alsace
contemporaine*

Anaïs KOEHLER, 3^e année de doctorat

Mots-clés : sociolinguistique, contact de langues, dialectologie

La contribution proposée a son focus sur un type d'occurrences relevé dans notre corpus, regroupé dans la catégorie de travail provisoire « code-switching » (CS) et analysé dans notre thèse portant sur « L'impact du français sur les parlers dialectaux de l'Alsace contemporaine ». Il s'agit d'occurrences en français apparaissant en co-texte alsacien, dont nous avons conservé les éléments syntagmatiques les plus longs pour l'analyse.

Les propriétés et fonctions des CS, qui sont classiquement étudiés sous un angle lexicogrammatical et donnent lieu à une liste, peuvent se combiner quand ils sont analysés sous un angle sociolinguistique et pragmatique. La complexité des combinatoires, si tant est que l'on parvienne à isoler les propriétés et à repérer toutes les combinaisons de caractéristiques de CS, entraîne une difficulté de classification/typologisation des CS : la raison de la survenue des CS peut ne pas apparaître, ou bien ceux-ci peuvent être aussi bien mono- que polyfonctionnels.

S'ajoute à cette difficulté de détermination la grande porosité du continuum dialecte alsacien – français qui amène à remettre en question la dénomination « code-switching » ou « alternance de code ». La prise en compte de l'usage et des pratiques sociales soulève des questions de fond sur le mode d'expression des locuteurs dialectophones : le CS peut-il dans tous les cas être désigné comme « alternance de code », ce qui suppose, chez le locuteur, deux « codes » séparés et un choix intentionnel et/ou conscient ?

La contribution est donc plutôt méthodologique : elle expose comment une difficulté pratique (détermination de la/des cause/s des CS) s'ajoute à un problème théorique (CS est-il le bon concept pour rendre compte de ce que font les dialectophones alsaciens ?) au milieu du travail doctoral en cours.

Références bibliographiques :

Alby, Sophie. 2013. « Alternances et mélanges codiques ». In *Sociolinguistique du contact. Dictionnaire des termes et concepts*, 43-70. Lyon : Éditions de l'École Normale Supérieure de Lyon.

Erhart, Pascale. 2018. « Les effets de la frontière sur les pratiques linguistiques dans le Rhin supérieur », *Les Cahiers du GEPE*, N°9/2017. *Migration(s) et langues ; langues et espace(s)*, URL : <http://cahiersdugepe.fr/index.php?id=3100>

Huck, Dominique, Bothorel-Witz Arlette et Geiger-Jaillet Anemone. 2007. « L'Alsace et ses Langues. Éléments d'une description sociolinguistique en zone frontalière ». In *Aspects of Multilingualism in European Border Regions: Insights and Views from Alsace, Eastern Macedonia and Thrace, Lublin Voivodeship and South Tyrol*, édité par Abel Andrea, Stuffleser Mathias et Voltmer Leonhard, 13-100. Bolzano : EURAC Research, pp.14-100

Huck, Dominique et Erhart Pascale. 2019. « Das Elsass ». In *Handbuch des Deutschen in West- und Mitteleuropa: Sprachminderheiten und Mehrsprachigkeitskonstellationen*, 155-192. Tübingen: Narr Francke Attempto.

Riehl, Claudia Maria. 2009. *Sprachkontaktforschung: Eine Einführung*. 2e éd. Tübingen : Gunter Narr Verlag.

———. 2014. *Mehrsprachigkeit: Eine Einführung*. Darmstadt: Wissenschaftliche Buchgesellschaft.

Études précédentes sur le code-switching en Alsace :

- Dorffer, Clément. 2019. « Dynamiques et stratégies du code-switching entre l'alsacien et le français ». Strasbourg : Université de Strasbourg.
- Fischer, Anne. 1979. « Les bilingues et le code-switching : Etude de l'alternance français - dialecte à Sarrebourg ». Strasbourg : Université de Strasbourg.
- Gardner-Chloros, Pénélope. 1985a. « Choix et alternance des langues à Strasbourg ». Thèse de doctorat, Strasbourg : Université Louis Pasteur.
- . 1985b. « Le code-switching à Strasbourg ». In *Le français en Alsace : actes du colloque de Mulhouse* (17-19 novembre 1983), 51-60. Genève : Editions Slatkine.
- . 2009. *Code-switching*. Cambridge: Cambridge University Press.
- Maurer, Johanna. 2002. « Elsässisch und Französisch : die Funktion ihrer Alternanz im Diskurs ». Kiel: Westensee-Verlag.

L'éthique dans la recherche sur les réseaux sociaux : Le consentement éclairé sur

Instagram

Erin McINERNEY, 1^e année de doctorat

Mots-clés : réseaux sociaux, Instagram, éthique, Web 2.0, identité numérique, sociolinguistique

L'éthique fait partie intégrante de la recherche universitaire. Elle constitue le fondement d'un ensemble de données productives qui peuvent être interprétées et, plus tard, utilisées par d'autres chercheurs. En parallèle, la divulgation des méthodes utilisées par les chercheurs pour obtenir leurs données est primordiale. Mais que faut-il faire quand il existe peu de méthodologies réputées pour une nouvelle recherche ? Alors que les nouveaux médias ont commencé à inonder la société, se mêlant presque parfaitement aux habitudes quotidiennes des jeunes générations, ils ont attiré l'attention des chercheurs dans de multiples domaines. En consultant les données librement accessibles sur les réseaux sociaux, quelles limites les chercheurs doivent-ils s'imposer ? Dans ce tournant vers la recherche sur les nouveaux médias, les chercheurs doivent, à leur tour, remettre en question les méthodes traditionnellement utilisées pour recueillir les données, en s'appuyant sur la posture avancée par boyd (2010) : « ce n'est pas parce que les données sont accessibles que leur utilisation est éthique. »

Pour élaborer cette problématique, nous citerons quelques principes clés de l'éthique dans la recherche pour ensuite voir si ou comment elles se transforment dans un contexte digital. Cette communication a pour but de mener une réflexion sur les méthodes proposées dans la collection et traitement des données sur les réseaux sociaux, ainsi que les possibilités de recherche qui s'offrent aux linguistes sur la plateforme Instagram.

Références bibliographiques :

- Bayard, C. (2018). Les mères célèbres sur Instagram: Ce que nous révèlent leurs mises en scène de l'allaitement. *Enfances, Familles, Generations*, 31, 187–206.
- Beninger, K. (2016). *Social Media Users' Views on the Ethics of Social Media Research in Sloan*, L., & Quan-Haase, A. *The SAGE Handbook of Social Media Research Methods*. SAGE Publications Ltd. <https://doi.org/10.4135/9781473983847>
- Blackwood, R. (2019). Language, images, and Paris Orly airport on Instagram: Multilingual approaches to identity and self-representation on social media. *International Journal of Multilingualism*, 16(1), 7–24. <https://doi.org/10.1080/14790718.2018.1500257>
- boyd, d. (2010). *Privacy and Publicity in the Context of Big Data*. Raleigh, North Carolina.
- Morelli, P. *Réseaux sociaux et éthique*. Université de Lorraine, France. <https://www.maisons-pour-la-science.org/>
- Moreno, M. A., Goniou, N., Moreno, P. S., & Diekema, D. (2013). Ethics of Social Media Research: Common Concerns and Practical Considerations. *Cyberpsychology, Behavior, and Social Networking*, 16(9), 708–713. <https://doi.org/10.1089/cyber.2012.0334>
- Lacaze, G. (2020). Les matérialités discursives et les pratiques citationnelles du discours rapporté numérique sur Instagram. *SHS Web of Conferences*, 78, 01004. <https://doi.org/10.1051/shsconf/20207801004>
- Piazzesi, C., & Lavoie Mongrain, C. (2020). Selfies de femmes, négociation normative et production de culture visuelle sur Instagram et Facebook. *Recherches féministes*, 33(1), 135. <https://doi.org/10.7202/1071246ar>

Voilà en français parlé contemporain : du déictique au marqueur de discours

Amandine Robert, 2^e année de doctorat

Nous nous intéresserons dans cette contribution aux principaux emplois oraux de *voilà*, mot en pleine expansion en français parlé contemporain. Comme le rappelle son étymologie – impératif du verbe ‘voir’ et déictique ‘là’ – cette unité est caractérisée par une forte dimension monstrative. Nous établirons deux catégories principales selon le domaine où se construit la référence : un *voilà* déictique lorsqu’il demeure possible d’identifier un référent dans le monde extralinguistique ou le co-texte (deixis au sens large) ; un *voilà* relevant du marqueur de discours, dont la portée est désormais l’énoncé, qu’il s’agisse du tour de parole d’un seul et même locuteur ou d’un énoncé co-construit. La présentation des usages discursifs de *voilà* sera accompagnée d’exemples tirés d’un corpus audio et vidéo constitué au cours des derniers mois (émissions radiophoniques / télévisuelles ; conseils de classe en lycée).

Cet intérêt pour *voilà* s’inscrit dans l’étude contrastive de certains marqueurs de discours de l’allemand et du français que je mène dans le cadre du doctorat. L’approche est principalement pragmatique et traductologique : il s’agit de déterminer de manière fine le rôle que joue le marqueur selon la situation d’énonciation et de rechercher dans la langue cible des équivalents aux fonctions pragmatiques similaires.