



CALL FOR PAPERS

COSEDI

SPECIALISED GENRE-BASED CORPORA: CHARACTERISATION, METHODS AND PEDAGOGICAL APPLICATIONS

**Université Grenoble Alpes
6-8 December 2023**

**Website: <https://cosedi.sciencesconf.org/>
Email: cosedi@univ-grenoble-alpes.fr**

COSEDI (*Corpus de genres spécialisés : caractérisation, méthodes et applications didactiques*) follows previous conferences held at the University of Grenoble Alpes (UGA) in the fields of Languages for Specific Purposes (LOSP), Lexicology, Terminology and Translation (LTT) as well as Linguistic resources and pedagogical applications in specialised languages (RADELAS). In 2023, COSEDI will take place in Grenoble from December 6th to December 8th.

In France, the research carried out in the late 90s (Habert et al., 1997) paved the way to major developments in corpus linguistics, which has since become a well-established discipline (Frérot & Pecman, 2021). Both the applications of corpora -e.g. translation, terminology, language teaching- and the research on building and analysing corpora testify to the prominence of corpus linguistics as shown by the body of literature in the field.

The notion of genres is of crucial importance when one needs to build and classify corpora as well as to understand discourses produced by specialised domains. Mastering the genres frequently used in a field means being recognised as a legitimate member of a specialised community (Swales 1990; Bhatia 2004). Genres, having a typical internal structure and, therefore, being conventionalised forms through which discursive practices exercise their performative form, represent a culturally recognised message (Biber et al., 2007). Although there are a number of genre studies, they have not yet produced any tools for empirical analyses that can be applied on a large scale across a variety of texts (*ibid.*). Hence, researchers lack knowledge about “the general patterns of discourse organization” across a representative text sample (*ibid.*).

Many researchers in language for specific purposes (LSP) rely on corpora based on specialised genres (Millot, 2018; Dury & Picton, 2009; Couterut, 2016; Parodi, 2008; Venegas, 2008; Calvi, 2010; De Beni & Hourani-Martin, 2020), thus demonstrating the interest of the intersection of the notions of corpus and specialised genre for this type

of study. The goal of this conference is to confront these two theoretical frameworks to better understand how they can mutually enrich each other. As Handford (2010, p. 256) points out:

This complementary combination of corpus and genre is a logical and desirable development in discourse analysis: corpora have much to say about language, but they can be lacking in contextual interpretability; genres are intrinsically contextual entities, but their linguistic features may be underexposed.

Topic 1: Definition and Characterisation of Specialised Genres-Based Corpora

The notion of a specialised genre has kindled many a thought and has been approached in very different ways, of which we can quote the new rhetoric (Freedman & Medway, 1996; Bazerman, 1994, 2008), applied linguistics (Swales, 2004; Bhatia, 2004), the semio-discursive perspective (Charaudeau, 2004), Sydney School (Halliday & Hasan, 1989; Martin & Rose, 2008), the Germanic communicative perspective (Heinemann, 2000) or discourse analysis (van Dijk, 2008).

The conference will focus more specifically on the generic approach for the definition and the characterisation of specialised corpora using tools provided by corpus linguistics. Indeed, the generic approach has proven particularly successful as a conceptual and methodological tool for the description of languages for specific purposes (Garcia Izquierdo, 2011). This approach is useful in a multidimensional conceptualisation of the texts produced by professional communities (Bolivar & Parodi, 2015) and, therefore, useful to a better understanding of dynamics of the aforementioned communities. Bathia (2014) furthermore highlights the benefit of this approach to explain production and reception processes.

It is assumed that specialised genres constitute sets of linguistic and textual features that can be identified from representative corpora that present prototypical regularities characterising a specific genre on a higher level of abstraction (Parodi, 2009).

Specialised corpora are of great help to shed light in the classification of genres in a given sector (Calvi, 2010) and, therefore, allow a differentiation of families of genres, macro-genres, genres and sub-genres (Ezpeleta Piorno & Gamero Pérez, 2004). However, the question of the parameters to be taken into account in defining genres continues to be a difficult task for corpus linguistics. This is why we aim to bring answers in this domain.

In this axis, we will also discuss the contributions of corpus linguistics to the definition of FASP (*Fiction à substrat professionnel* or profession-based fiction) as a specialised genre. For the past 20 years, FASP has had a significant cultural and linguistic documentary value, due to its focus on the professional substratum. It provides insight into aspects of the professional setting, including the specialised encyclopaedia of the fields concerned (Van der Yeught, 2016) and familiarity with the specialised lexicon (Petit, 1999). A wide variety of FASP sub-genres have been identified in the literature: legal, medical, journalistic, academic, scientific, ecological, artistic... (Isani, 2010). A central question concerns the use of specialised corpora to characterise the different types of FASP sub-genres.

Topic 2: Methodological Issues Related to Collecting and Analysing Specialised-Genre Corpora

Collecting specialised genre-based corpora entails a number of well-documented methodological issues. The smaller size of specialised corpora allows for more contextualisation (Koester, 2010), especially since they are composed of full texts (Wozniak, 2019, p.122). The links between quantitative and qualitative analyses are, therefore, questioned, in particular how to take into account the extralinguistic characteristics of the specialised domains concerned and their integration in corpus analysis (Nesi, 2013).

The issue of the representativeness of the selected texts and the possible biases in the construction of the corpus is also central (Nelson, 2010, p. 57; Cabré, 2007). For example, studies on scientific discourse, such as those conducted in the ANT Scientext or TermITH projects for French, must integrate a diverse panel of disciplines and types of scientific productions (Tutin & Grossmann, 2014; Jacques & Tutin, 2018). The notion of representativeness is all the more essential when building diachronic corpora aiming to represent the evolution of a specialised variety (Dury, 2004; Millot, 2018). Finally, the question of data access is fundamental, especially for the constitution of corpora concerning professional genres (Wozniak, 2019) or oral corpora (Divoux, 2021; Surcouf, 2021), which entail constraints in terms of access to information, confidentiality and ethics.

In this context, proposals will focus on the definition of relevant methodological criteria for the construction or analysis of specialised genre-based corpora. Their analysis will also be explored and proposals focusing on tools and methods adapted to this type of corpus will be welcome. The issue of non-ecological corpora (FASP, role-playing games, simulations) as a way of accessing specialised genres will also be addressed.

Topic 3: Using Specialised Genre-Based Corpora for Translation and Terminology Purposes

In specialised translation, previous research has indicated that corpora, and comparable corpora in particular, are a useful tool in the decision-making process and the problem-solving process involving terminology and phraseology (Kübler, 2003; Maia, 2003; Varantola, 2003; Marco & van Lawick, 2009). Over the past few years, the availability of large online translation corpora and machine translation tools has revolutionized the translation process as a whole. From the pedagogical perspective, it has raised new issues and it poses a new challenge to language teachers and translator trainers. To that respect, combining the use of comparable corpora and translation corpora through post-editing in the translation classroom has been a new and innovative area of research (Kübler et al., 2022).

Terminology has benefited greatly from the corpus-based approach. The research carried out on building terminological resources and extracting knowledge from corpora has contributed to establishing corpus-based terminology as a discipline with its own methods and principles (Bourigault et al., 2001). Much of the linguistic research has focused on identifying «*knowledge-rich contexts*» (hence KRCs, Meyer, 2001) and relation markers across specialised domains and text genres in monolingual or comparable corpora (Marshman, L'Homme, 2008; Picton, 2009; Lefeuve, 2017; Ledouble, 2021). Recent studies have shown how the use of translation corpora could help identify both KRCs and relations markers (Condamines et al., 2021).

Proposals on the use of corpora and tools for translation and terminology purposes in the light of the current research are particularly welcome. Contributors are invited to explore issues on specialised knowledge domains and genres within a monolingual or

bilingual perspective. Proposals may also include any aspect related to developing corpus-based tools for translation and terminology purposes.

Topic 4: Pedagogical Applications of Specialised Genre-Based Corpora

As mentioned above, the use of corpora makes it possible to distinguish various representative features of the genres in question and to provide learners with appropriate instruction, using the evidence-based approach (Hyland, 2006). The joint use of corpora and genres is often documented in teaching academic writing (cf. Tribble, 2002; Swales & Lee, 2006), the latter occasionally relying on the comparison of output between learners and experts (Yan & Jacques, 2022).

As far as the use of corpora in language teaching and learning is concerned, indirect and direct approaches, often referred to as data-driven learning, co-exist (Johns, 1991; Gollin-Kies et al., 2015). From a more pragmatic perspective, genre-based tasks can be divided into three categories: rhetorical consciousness-raising, text production and process activities (Hyon, 2018).

The use of corpora in general, and of specialised genres in particular, in LSP courses offers several advantages: relevance to professional reality, development of academic literacy, improvement of transversal skills, positive influence on learners' attitudes, possibility to work on linguistic aspects specific to the chosen genres, etc. (Flowerdew, 2009; Landure & Boulton, 2010).

We welcome proposals on any pedagogical aspect of the use of specialised genre-based corpora.

References

- Baker, P. (2006). *Using corpora in discourse analysis*. London, UK: Continuum.
- Bazerman, C. (1994). Systems of genres and the enhancement of social intentions. In A. Freedman & P. Medway (Eds.), *Genre and new rhetoric* (pp. 79-101). London, UK: Taylor & Francis.
- Bazerman, C. (Ed.). (2008). *Handbook of research on writing. History, society, school, individual, text*. New York, NY: Erlbaum.
- Bhatia, V. (2004). *World of written discourse: A genre-based view*. London, UK: Continuum.
- Bhatia, V. & Bremner, S. (Eds.). 2014. *The Routledge handbook of language and professional communication*. Oxon, UK; New York, NY: Routledge
- Biber, D., Connor, U., & Upton, T. (2007). *Discourse on the move. Using corpus analysis to describe discourse structure*. Amsterdam, the Netherlands/ Philadelphia, PA: John Benjamins Publishing Company.
- Bolívar, A., & Parodi, G. (2015). Academic and professional discourse. In M. Lacorte (Ed.), *The Routledge handbook of Hispanic applied linguistics* (pp. 459–476). New York, NY: Routledge.
- Bourigault, D., Jacquemin, C., & L'Homme, M.-C. (Eds.) (2001). *Recent advances in computational terminology*. Amsterdam, the Netherlands/ Philadelphia, PA: John Benjamins Publishing Company.
- Cabré, M. T. (2007) Constituer un corpus de textes de spécialité. *Cahier du CIEL*, 20, 37–56.
- Calvi, M.V. (2010). « Los géneros discursivos en la lengua del turismo: una propuesta de clasificación ». *Ibérica* 19, 9–31.
- Charaudeau, P. (2004). La problemática de los géneros. De la situación a la construcción textual. *Revista Signos*, 37(56), 23–39.

- Condamines, A., Escoubas, M.-P., & Federzoni, S. (2021). Apport de la traduction dans l'analyse des marqueurs de relations conceptuelles. Une étude en corpus aligné français-italien. In C. Frérot & M. Pecman (Eds.), *Des corpus numériques à l'analyse linguistique en langues de spécialité* (pp. 313–336). Grenoble, France: UGA Editions.
- Coutherut, M. (2016). *Les textes procéduraux en anglais : création d'une échelle de prototypicalité* (Thèse de doctorat). Sorbonne Paris Cité, Paris.
- Divoux, A. (2021). Étudier les interactions verbales en milieu professionnel : aspects méthodologiques du recueil de données en réunion de travail. In C. Frérot & M. Pecman (Eds.), *Des corpus numériques à l'analyse linguistique en langues de spécialité* (pp. 87–106). Grenoble, France: UGA Editions.
- Dury, P. (2004). Building a bilingual diachronic corpus of ecology: The long road to completion. *ICAME Journal*, xxviii, 5–16.
- Dury, P., & Picton, A. (2009). Terminologie et diachronie : vers une réconciliation théorique et méthodologique? *Revue française de linguistique appliquée*, 14(2), 31–41.
- Ezpeleta Piorno, P., & Gamero Pérez, S. (2004). « Los géneros técnicos y la investigación basada en corpus: el proyecto GENTT ». In R. Gaser, C. Guirado, & J. Rey (Eds.), *Insights into scientific and technical translation*, 147–156.
- Flowerdew, L. (2009). Applying corpus linguistics to pedagogy: A critical evaluation. *International Journal of Corpus Linguistics*, 14(3), 393–417.
- Freedman, A., & Medway, P. (Eds.). (1994). *Genre and the new rhetoric*. London, UK: Taylor & Francis.
- Frérot, C., & Pecman, M. (Eds.). (2021). *Des corpus numériques à l'analyse linguistique en langues de spécialité*. Grenoble, France: UGA Éditions.
- García Izquierdo, I. (2007). « Los géneros y las lenguas de especialidad ». In E. Alcaraz Varó, J. Mateo Martínez, & E. Yus Ramos (Eds.), *Las lenguas profesionales y académicas* (pp. 119–125).
- Gollin-Kies, S., Hall, D., & Moore, S. (2015). *Language for specific purposes*. Basingstoke, UK: Palgrave Macmillan.
- Habert, B., Nazarenko, A., & Salem, A. (1997). *Les linguistiques de corpus*. Paris, France: Armand Colin.
- Halliday, M., & Hasan, R. (1989). *Language, context and text aspects of language in a social – semiotic perspective*. Hong Kong: Oxford.
- Handford, M. (2010). What can a corpus tell us about specialist genres. In A. O'Keeffe & M. McCarthy (Eds.), *The Routledge handbook of corpus linguistics* (pp. 255-269). London, UK: Routledge.
- Heinemann, W. (2000). Textsorten. Zur Diskussion um Basisklassen des Kommunizierens. Rückschau und Ausblick. In K. Adamzik (Ed.), *Textsorten* (pp. 2–29). Tübingen, Germany: Stauffenburg Verlag Brigitte Narr GmbH.
- Hyland, K. (2006). *English for academic purposes*. London, UK: Routledge.
- Hyon, S. (2018). *Introducing genre and English for specific purposes* [Kindle edition]. Oxon, UK: Routledge.
- Isani, S. (2010). « Avant-propos ». *ILCEA* [En ligne], 12. URL: <http://journals.openedition.org/ilcea/443>.
- Jacques, M. P., & Tutin, A. (2018). *Lexique transversal et formules discursives des sciences humaines*. London, UK: ISTE Group.
- Johns, T. (1991). Should you be persuaded: Two examples of data-driven learning. *Classroom Concordancing. English Language Research Journal*, 4, 1–16.
- Koester, A. (2010). Building small specialised corpora. In A. O'Keeffe & M. McCarthy (Eds.), *The Routledge handbook of corpus linguistics* (pp. 66–79). London, UK: Routledge.

- Kübler, N. (2003). Corpora and LSP translation. In F. Zanettin, S. Bernardini, & D. Stewart (Eds.), *Corpora in translator education*. Manchester, UK: St Jerome.
- Kübler, N., Mestivier, A., & Pecman, M. (2022). Using comparable corpora for translating and post-editing complex noun phrases in specialised texts: Insights from English-to-French in specialised translation. In S. Granger & M-A. Lefer (Eds.), *Extending the scope of corpus-based translation studies* (pp. 237–266). London, UK: Bloomsbury Publishing.
- Landure, C., & Boulton, A. (2010). Corpus et autocorrection pour l'apprentissage des langues. *ASp*, 57, 11–30.
- Ledouble, H. (2021). Contextes et connaissances dans les discours de vulgarisation du scientifique : dynamique définitoires et problématiques cognitives. In C. Frérot & M. Pecman (Eds.), *Des corpus numériques à l'analyse linguistique en langues de spécialité* (pp. 285–312). Grenoble, France: UGA Editions.
- Lefeuvre, L. (2017). Analyse des marqueurs de relations conceptuelles en corpus spécialisé : Recensement, Évaluation en fonction du domaine et du genre textuel (Thèse de doctorat). Université Toulouse Jean Jaurès, Toulouse.
- Maia, B. (2003). "Some languages are more equal than others": Training translators in terminology and information retrieval using comparable and parallel corpora. In F. Zanettin, S. Bernardini & D. Stewart (Eds.), *Corpora in translator education* (pp. 47–58). Manchester, UK: St Jerome.
- Marco, I., & van Lawick, H. (2009). Using corpora and retrieval software. In A. Beeby, P. Sánchez-Gijón, & I. P. Rodríguez (Eds.), *Corpus use and translation: Corpus use for learning to translate and learning corpus use to translate* (pp. 9–28). Amsterdam, the Netherlands/ Philadelphia, PA: John Benjamins Publishing Company.
- Martin, J. R., & Rose, D. (2008). *Genre relations: Mapping culture*. London, France: Equinox.
- Marshman, E., & L'Homme, M.-C. (2008). Portabilité des marqueurs de la relation causale : étude sur deux corpus spécialisés. In F. Maniez, P. Dury, N. Arlin, & C. Rougemont (Eds.), *Corpus et dictionnaires de langues de spécialité* (pp. 87–110). Grenoble, France: Presses universitaires de Grenoble.
- Meyer, I. (2001). Extracting knowledge-rich contexts for terminography: A conceptual and methodological framework. In D. Bourigault, C. Jacquemin, & M.C. L'Homme (Eds.), *Recent advances in computational terminology* (pp. 279–302). Amsterdam, the Netherlands/ Philadelphia, PA: John Benjamins Publishing Company.
- Millot, P. (2018). L'évolution terminologique des sciences de l'information : étude sur corpus diachronique d'un nouveau domaine spécialisé. *Textes et contextes* [online], 13.2, mis en ligne le 05 juillet 2019, consulté le 16 janvier 2023. URL: <http://preo.u-bourgogne.fr/textesetcontextes/index.php?id=2318>
- Nelson, M. (2010). Building a written corpus: what are the basics? In A. O'Keeffe & M. McCarthy (Eds.), *The Routledge handbook of corpus linguistics* (pp. 53–65). London, UK: Routledge.
- Nesi, H. (2013). ESP and corpus studies. In B. Paltridge & S. Starfield (Eds.), *The handbook of English for specific purposes* (pp. 407–426). Oxford, UK: Wiley-Blackwell.
- Parodi, G. (2009). El corpus académico y profesional del español PUCV-2006: semejanzas y diferencias entre los géneros académicos y profesionales. *Estudios filológicos*, 44, 123–147.
- Petit, M. (1999). « La fiction à substrat professionnel : une autre voie d'accès à l'anglais de spécialité », *Asp* 23-26.
- Picton, A. (2009). Diachronie en langue de spécialité, définition d'une méthode linguistique outillée pour repérer l'évolution des connaissances en corpus : un exemple appliqué au domaine spatial (Thèse de doctorat). Université Toulouse 2, Toulouse.
- Surcouf, C. (2021). Le français oral quotidien, un objectif spécifique en FLE ? Retour sur les défis de la création d'un corpus de français parlé annoté à visée pédagogique. In C. Frérot

- Cécile & M. Pecman (Eds.), *Des corpus numériques à l'analyse linguistique en langues de spécialité* (pp. 87–106). Grenoble, France: UGA Éditions.
- Swales, J.M. (1990). *Genre analysis*. Cambridge, UK: Cambridge University Press.
- Swales, J.M. (2004). *Research genres: Exploration and applications*. Cambridge, UK: Cambridge University Press.
- Swales, J.M., & Lee, D. (2006). A corpus-based EAP course for NNS doctoral students: Moving from available specialized corpora to self-complied corpora. *English for Specific Purposes*, 25, 56–75.
- Tribble, C. (2002). Corpora and corpus analysis: New windows on academic writing. In J. Flowerdew (Ed.), *Academic discourse* (pp. 131–149). Harlow, UK: Pearson Education Limited.
- Tutin, A., & Grossmann, F. (Eds.). (2014). *L'écrit scientifique: du lexique au discours*. Rennes, France: Presses universitaires de Rennes.
- Van der Yeught, M. (2016). A proposal to establish epistemological foundations for the study of specialised languages. *ASp*, 69, 41–63. <https://doi.org/10.4000/asp.4788>
- van Dijk, T. (2008). *Discourse and context. A sociocognitive approach*. Cambridge, UK: Cambridge University Press.
- Varantola, K. (2003). Translators and Disposable Corpora. In F. Zanettin, S. Bernardini, & D. Stewart (Eds.), *Corpora in translator education* (pp. 55–70). Manchester, UK: St. Jerome Publishing.
- Venegas, R. (2008). « La densidad informacional en registros diversos: Aproximación computacional utilizando el 'Manchador de Textos' ». *Revista de lingüística teórica y aplicada (RLA)*, 46(2), 53–79.
- Wozniak, S. (2019). *Approche ethnographique des langues spécialisées professionnelles. Aspects linguistiques et cultures des discours spécialisés*. Bern, Switzerland: Peter Lang.
- Yan, R., & Jacques, M. P. (2022). Les routines verbales en *comme* : diagnostic des besoins langagiers du FLE par une approche comparative. In M. Luodonpää-Manni, F. Grossmann, & A. Tutin (Eds.), *Les routines discursives dans le discours scientifique oral et écrit* (pp. 213–232). Grenoble, France: UGA Éditions.

Plenary Speakers

Mojca Pecman, Université Paris Cité (CLILLAC-ARP, UFR EILA)

Paolo Frassi, Université de Vérone

John Flowerdew, Professeur invité, Université de Lancastre, anciennement

Professeur à l'Université municipale de Hong Kong



Important dates

Deadline for proposals : **25 Juin 2023**

Evaluation of the proposals by the scientific committee : June-September 2023

Information to authors : September 2023

Confirmation of participation by authors : early October 2023

The presentation of contributions in French or in English will not be longer than 3 pages (excluding references and figures). **Anonymous proposals will be submitted via ScienceConf** for evaluation by two reviewers.