Winter School on
Corpus Data for the Analysis of Discourse, Interactions and Arguments (February 7-9, 2023)

In the context of *LAND: Doctoral Programme in Applied Linguistics: Managing Languages, Arguments and Narratives in the Datafied Society*

Abstracts
of student presentations for
Data Sessions

Elisa Angiolini (UNINE): *Subdiscussions in children-adult argumentative interactions*
In the domain of children's argumentation, scholars have pointed out that children can actively participate in argumentative discussions, meeting certain circumstances (Bubikova-Moan & Sandvik, 2022; Greco & Perret-Clermont, 2022). My dissertation focuses on the under-investigated phenomenon of young children opening subdiscussions, which is when, within an ongoing argumentative discussion, some aspect of a party's utterance is put into question and rediscussed (cf. Schär, 2021). After describing the aims of my project, I will focus on the operationalization of my research questions, then I will present how I collected my data, how I transcribed it, and the argumentative analysis I have been carrying out. By doing this, I will expose some challenges and doubts that I have been encountering in this process.

References
Bubikova-Moan, J., & Sandvik, M. (2022). Argumentation in early childhood: A systematic review. Human Development.
Greco, S., & Perret-Clermont, A. N. (2022). Young Children's Argumentative Contributions. In J. Fahnestock & R.A. Harris (eds.). The Routledge Handbook of Language and Persuasion (pp. 457-474). Routledge.
Schär, R. G. (2021). An argumentative analysis of the emergence of issues in adult-children discussions. John Benjamins.

Elena Battaglia (USI): Evidentiality in Italian talk-in-interaction. Incremental and interactive categorization of sources of knowledge.
As a grammatical and functional category, evidentiality pools the array of strategies - verbal and non-verbal - whereby speakers categorize their sources of knowledge for an utterance.
At the crossroads of research on evidentiality and grammar in interaction, my PhD work investigates the distribution and combination of grammatical, lexical and textual strategies in discourse, and finds that the verbal encoding of sources knowledge is an emergent and locally negotiated endeavor, finely tuned to the temporality, sequentially and intersubjectivity of spoken language.
In this talk, I will present some theoretical tools, namely evidential frames and constructions, methodological issues, and ongoing analyses. My data include informal conversations among friends from the TIGR (6h) and Kiparla (16h) corpora of spoken Italian. TIGR data are currently being annotated, building on the Infinita (FNS 192771, USI 2020-2024) project's scheme, in view of a quantitative appraisal of the distribution of evidential strategies across turn constructional units and sequential positions. Kiparla data form the basis for a complementary collection of cases documenting instances of "incremental and interactive" categorizations of sources of knowledge, in view of a qualitative analysis of their pragmatic and interactional functions.

Barcella Deborah (USI): *The strategic communication of the World Wide Web: a qualitative analysis of WWW's communication through the CERN archive*
In 1989, some people at CERN developed the World Wide Web, an Internet network information system, to enhance communication within the laboratory (Berners-Lee, 2000). People in the field of marketing have looked at the Web as a powerful tool for advertising, but no one has ever investigated how this technology was promoted to so many people. My dissertation aims to fill a gap in the literature on the World Wide Web by shifting the focus from the Web as a marketing channel to the Web as a product that was communicated strategically through marketing activities. Three papers will serve as the foundation for my Ph.D. thesis: 1) the Web's brand identity; 2) its promotional activities; and 3) its market dynamics and why it is free. To investigate this topic, I am conducting a content thematic analysis on the archives of CERN, to which my research team has exclusive access and whose materials are extremely heterogeneous (handwritten, printed, and visual material).
References
Berners-Lee, T. (2000). Weaving the Web: The Original Design and Ultimate Destiny of the World Wide Web (1st edition). Harper Business.

José Alfonso Lomelí Hernández (USI): *Crossing data between corpora: from press conferences to twitter*
The presentation discusses the challenges of retrieving and analyzing data from two different corpora. The first corpus is made of press conferences held by politicians during the Covid-19 pandemic, while the second corpus is made of tweets posted by users concerning the press conferences of the first corpus. First, the presentation discusses the technical challenges of gathering the data of the second corpus, and the difficulties of analyzing the tweets of users commenting on the press conferences. Later, the presentation moves to discuss the research questions of the study and the preliminary results of the analysis.

Dario D'Agostino (ZHAW): *Heuristics towards an analysis of deictic practices in digital discourse*
From a Foucauldian perspective, (societal) discourse is organized, controlled, selected, and redistributed by a set of procedures. By the principle of the Commentary, discourse is ordered by a gradation among "text": there are primary, canonical texts which are perpetually recited, repeated, and varied by secondary, commentary texts.
In digital discourse, hyperlinking (as a deictic practice of pointing) can be analyzed in the framework of the Commentary principle to gain structural insights into discourse. Methodologically, approaches from corpus-linguistics and visual linguistics allow a data-oriented analysis: With corpus analysis the actual linguistic realization of the hyperlinks can be retrieved and by doing so, deictic patterns can be detected. With network analysis, a deictic network can be generated and used to explore the Commentary principle. In the presentation, heuristics are discussed which cover the process of data selection, retrieval, processing, and analysis. The presentation includes both, solutions and open problems.

Giulia D'Agostino (USI): *Pretty little lies. Uncovering dialectical and argumentative patterns emerging from irregular or unsatisfactory answers in the financial dialogue*
In the story we are going to tell we have two main actors: listed companies and investors – or better, financial analysts representing investors' interests. We then have a main setting for their natural interaction: an institutionalized yet relatively informal environment where they can disclose opinions, thoughts, visions, and reasonings – about the company, its management, the course of action that it undertook or is going to endorse.
In such a context, strategic communication and argumentative expertise are key: companies need to persuade investors of the soundness of their financial choices, whereas capital holders demand transparency to better evaluate the advisability of investment.
But what about situations when managers are aware they better hide some details, or analysts suspect there has been some concealing attempt? The cooperation between parts is somewhat disrupted, on a spectrum of overtness of such an irregularity. Strategies of dialogical evasiveness or cornering, from either side, are varied and to some extent attested. Taking cue from them and being aware of some structural obstacles –

this will lead us to focusing on overt cases only, of which some might be false positives – we will primarily progress towards:
• Listing cases of either irregularity
• Constructing a typology of both
• Underlining potential relationships between the two
• Tracing recurrent patterns along the downstream interaction when one such case is found.
The current session will therefore discuss approaches and emerging struggles in the analysis of data for (some of) the goals listed above.

Costanza Lucchini (USI): *Epistemics and evidentials in ECCs*
ECCs are teleconferences held by corporate leaders with financial analysts immediately following the quarterly earnings announcement. This highly-conventionalized activity type, akin to press conferences, represents a key event influencing investors' decisions and, in turn, affecting market outcomes. In this sense, access to information is a crucial asset both for companies and for the investing community; however, in this interaction, a clear *informative asymmetry* between managers and analysts persists.

Epistemic and *evidential* markers can be valuable indicators to map the knowledge management in ECCs, as well as to identify argumentative discourse relations. The first step of this study is the manual annotation of *epistemic frames* according to a newly developed annotation scheme. I will present my scheme and its implementation in INCEpTION, as well as the challenges that have already emerged, regarding, in particular, the operationalization of the scheme and the process of annotation.

Daniel de Oliveira Fernandes (University of Fribourg): *An experimental approach to the rhetorical effectiveness of insinuated ad hominem*
The research presented here is part of a project experimentally investigating the rhetorical advantages of implicit meaning in argumentation. In the experiments to be presented, we were interested in testing the difference between asserted and insinuated personal attacks in an abusive ad hominem setting (i.e., when an opponent refutes the proponent's argument by directly attacking the proponent personally).
For this purpose, 40 dialogues were created, each with a brief introductory context, a proponent presenting a standpoint, and the opponent refuting the proponent's standpoint by supporting their disagreement with either an insinuated attack, an assertive attack, or a neutral argument. To ensure that insinuations were interpreted as the asserted attacks but also left room for an alternative interpretation – which defines insinuations and their interests – pre-tests were conducted. We then pre-registered and ran our experiments to test our hypotheses. Contrary to our assumptions, (i) insinuations are not perceived as supporting a disagreement, (ii) nor leading to more agreement with the opponent. However, (iii) insinuations are perceived as more persuasive and (iv) more ethos-friendly (i.e., in terms of ability, benevolence, and integrity).

Chiara Mercury (USI): *Analysing conflicting frames and argumentation in the polylogical controversy surrounding fashion sustainability: preliminary findings and challenges*
Drawing from conflict resolution studies, the parties involved in controversy can be said to hold conflicting characterization frames, i.e. conflicting perspectives about themselves, their opponents, and their relationships (Shmueli, 2008). In this respect, an argumentative approach to the study of controversies can help to reconstruct the reasons underlying such conflicting frames. In order to determine the relationship between conflicting frames and argumentation in the context of the public controversy surrounding fashion sustainability, I consider a composite corpus including multi-genre documents issued by the different participants to the controversy. This public controversy represents an instance of argumentative polylogue (Aakhus and Lewiński, 2017). In my presentation, I will discuss the challenges that I am encountering in combining the communicative and linguistic levels of analysis, as well as in conducting a polylogical analysis that considers documents belonging to different genres.
Aakhus, M. & Lewiński, M. (2017). Advancing polylogical analysis of large-scale argumentation: disagreement management in the fracking controversy. Argumentation, 31, 179-207.

Shmueli, D. (2008). Framing in geographical analysis of environmental conflicts: Theory, methodology and three case studies. Geoforum, 39(6), 2048-2061.

Narjes Sheikh Asadi (USI): *Studies in Syntactic complexity, lexical sophistication, and argumentative patterns in a corpus of published journal articles written by native and non-native English writers*
This study investigates research articles' argumentative patterns focusing on native and non-native English writers. The attention to this issue is because there are far more L2 speakers of English than L1 English speakers, and these researchers have an essential need to publish their studies in prestigious journals. This study aims to compare the argumentative patterns, syntactic complexity, and lexical sophistication of research articles with different citation levels in Language and Literature journals with various quartiles between native and non-native scholars. In doing so, the introduction and discussion sections of research articles (RAs), which respectively provide the rationale for conducting the study and persuading the readers that the results of the study make sense and are comparable with previous ones, are extracted and classified in order to analyze the argumentative patterns based on a two-tier analysis; a modified system of moves and the analyzing standpoints in each move with pragma dialectics. Then, to analyze the corpus linguistically, two novel automated tools, The Automatic Assessment of Syntactic Sophistication and Complexity (TAASSC) and The Automatic Analysis of Lexical Sophistication (TAALES), are applied to calculate the paradigms' text scores for several classic and new indices of linguistic features, including syntactic complexity (the formal characteristics of syntax), and lexical sophistication (word frequency, range, bigram, and trigram frequency, academic language, and psycholinguistic word information). The research findings may assist academic writing, and authors in publishing more articles in prestigious journals that are embedded with better writing proficiency and deeper argumentation insight.

Ramy Younis (University of Fribourg): *Experimental investigations into the rhetorical effects of rephrase*
Empirical evidence from corpus data indicates that speakers rephrase frequently in argumentative discourse (Koszowy et al., 2022). Such findings call for an understanding of the rhetorical effects of this phenomenon: what persuasive advantages can be gained by rephrasing a contribution? In this talk, I discuss experimental evidence that we collected to shed light on this question.
Our set of experiments seeks to measure the effects of rephrase on the perceived persuasiveness of a message and the perceived trustworthiness of a speaker. Moreover, through our experiments, we seek to understand how participants perceive rephrase by asking them to evaluate the content similarity of segments that are connected through a rephrase relation. In the process, many questions deserving of attention have arisen: How can data collected through corpus analysis fuel experimental designs? What are the defining features of the rephrase relation? Are there clear boundaries between rephrase and inference relations?
References
Koszowy, M., Oswald, S., Budzynska, K., Konat, B., & Gygax, P. (2022). A Pragmatic Account of
Rephrase in Argumentation: Linguistic and Cognitive Evidence. Informal Logic, 42(1), 49–82.
https://doi.org/10.22329/il.v42i1.7212