

DARLING

Deniability in **AR**gumentation
and **LING**uistics

Kickoff workshop

Wednesday, April 29th 2026

Espace Guggi - MIS 08 0101

Programme

- 09:00-09:30** **Steve Oswald**
DARLING: Approaching denials at the interface of pragmatics and argumentation
- 09:30-10:00** **Annette Hautli-Janisz**
'Whoa, don't put words into my mouth!': How LLMs deny insinuations
- 10:00-10:30** **Diana Mazzarella**
'Networked deniability': How denials operate in digital communication
- 10:30-11:00** **Coffee break**
- 11:00-11:30** **Benedikt Pirker**
Experimental Linguistics and International Law: Connections to the DARLING Project - Bridging Pragmatics, Morality and Deniability
- 11:30-12:00** **Jan Albert van Laar**
To a dialectical account of clarification
- 12:00-13:30** **Lunch break**
- 13:30-14:00** **Madeleine Antille**
The impact of prosody on the comprehension of insinuation and plausibility of its denial
- 14:00-14:30** **Eda Sükan & Bernardo Rilla**
Analyzing Commitment Denial from a Multimodal Context
- 14:30-15:00** **Diane Liberatore**
Denial as a premiss of an accusation of bad faith: a means of defence and an indicative behaviour
- 15:00-15:30** **Coffee break**
- 15:30-16:00** **Bruna Paz Schmid**
Compiling DARLING's legal corpus: Denials in committee oversight hearings and where to find them
- 16:00-16:30** **Misha Müller**
Rethinking Plausible Deniability: Implicatures, Threats, and Legal Responsibility
- 16:30-17:00** **Lou Odermatt**
An argumentative approach to deniability: theoretical and experimental considerations

Abstracts

DARLING: Approaching denials at the interface of pragmatics and argumentation

Steve Oswald

University of Fribourg

In this talk I outline the rationale behind the DARLING project and present its main objectives. To better delineate the notions that we will grapple with, I first discuss pragmatic considerations about what goes into the notion of deniability, based on linguistic observations and extant literature on the topic. I then defend the relevance of an argumentative construal of denials and argue that opening up inquiry to the input of rhetoric and dialectics allows, in relevant ways, to tackle research questions related to their conversational effects and functions in verbal exchanges. I finish by a brief discussion of planned work, by summarising the project's main research questions and describing the two research strands (which recruit corpus and experimental methodologies) that the project will explore to answer them.

'Whoa, don't put words into my mouth!': How LLMs deny insinuations

Annette Hautli-Janisz

University of Passau

Despite the ubiquity of Large Language Models (LLMs) in current computational linguistic research, the phenomenon of deniability has only been touched on its fringes, for instance when discussing model sycophancy. This talk is the first to report work on the capabilities of LLMs to deny insinuations, based on data collected from the current state of the literature, experimental settings and real life. We identify whether (a) different models exhibit different denial patterns, (b) models can be steered towards different types of denials and (c) the generated denials are plausible. With a combination of quantitative and qualitative evaluation, we show that models are overall capable of denying their previous insinuations, but the nature of the responses varies between models, data instances and prompting technique.

‘Networked deniability’: How denials operate in digital communication

Diana Mazarella

University of Neuchâtel

This work examines how denials function in asynchronous, multimodal, and platform-distributed digital communication. I present the concept of networked deniability, a collectively produced form of deniability emerging from three interconnected mechanisms: network-enhanced context reconstruction, context proliferation, and epistemic destabilisation. Together, these mechanisms show how distributed actors collaboratively generate alternative interpretations of controversial communicative acts, reshaping interpretive disputes and complicating the attribution of responsibility in digital environments.

Experimental Linguistics and International Law: Connections to the DARLING Project - Bridging Pragmatics, Morality and Deniability

Benedikt Pirker

University of Fribourg & Liechtenstein Institut

The presentation explores the intersection of experimental linguistics and international law, focusing on how moral factors influence treaty interpretation—a topic directly relevant to the DARLING project’s investigation of deniability in argumentation. Drawing on experimental research that examines the role of moral biases in legal interpretation, the talk highlights how laypersons and experts alike are swayed by morally-valenced interpretations, often overriding explicit textual meaning. By employing a simplified typology of pragmatic interpretations (e.g., linguistic meaning, implicatures), the research demonstrates the psychological reality of “ordinary meaning” in legal contexts and underscores the importance of accounting for cognitive biases in argumentative discourse. The presentation will connect these findings to DARLING’s goals, particularly its focus on denials, accountability, and the rhetorical effects of pragmatic meaning.

To a dialectical account of clarification

Jan Albert van Laar

University of Groningen

In argumentative dialogue, proponents of arguments should be entitled to clarify their previous utterances in response to requests for clarification (Walton, 2007), but also in response to substantial criticisms (van Laar, 2010). The latter is necessary when proponents risk being or have been misunderstood. After all, it is impossible to be certain that one has adequately expressed propositional content. Furthermore, opponents may exploit the scope for interpretation by selecting interpretations that are not intended by, and unfavorable to, proponents. To address this, proponents should be entitled to advance a more precise reformulation. When proponents clarify arguments in this way, they uphold the initial propositional content and deny any content that fell within the scope of interpretation of the initial formulation but is excluded by the new reformulation. However, this entitlement can obviously be abused. What is presented as clarification may, on closer inspection, be an attempt to evade an existing commitment, whereby the proponent denies an actual propositional commitment that has at most a semblance of plausibility. How can we account for the entitlement to clarify and the potential to abuse it in a normative model of argumentation? First, we need to strike a balance when designing norms for critical discussion or persuasive dialogue (van Eemeren, 2010). At the same time we must acknowledge that abuse can often only be identified when a pattern of improper clarifications is displayed across various discussions. In this latter case, it may be helpful to examine norms for fair strategies, i.e. strategies that are balanced, transparent and tolerant, and that can be employed across various dialogues (van Laar and Krabbe, 2016). How can we respond to unfair strategies?

The impact of prosody on the comprehension of insinuation and plausibility of its denial

Madeleine Antille

University of Neuchâtel

This research investigates the impact of prosodic cues on the comprehension of insinuations and the plausibility of their subsequent denials. Sperber and Wilson's relevance theory describes the process of pragmatic inference in which contextual assumptions are used to derive the most relevant interpretation, leading to the comprehension of implicit meaning, and thus insinuations. Within this process, prosodic cues can act as guides that constrain interpretation toward intended meanings (Wilson and Wharton, 2006). From this, it follows that prosodic cues should facilitate the comprehension of insinuations while simultaneously reducing the plausibility of their denials, as they render salient the contextual assumptions underpinning the insinuated interpretation. Three experiments were conducted to test these predictions. Experiment 1 aimed to clarify the effect of RFR contour and a 500ms pause on the comprehension of insinuation and revealed these cues, when combined, allowed insinuations to be derived when the contextual cues were weak, but had no effect in the presence of strong contextual cues or individually. Experiments 2 and 3 investigated the impact of the same prosodic cues on the plausibility of denials, and the subsequent effects on the reputation of both speaker and target of the insinuation. In addition, these effects were analysed with respect to two types of denials, namely cancellations and disavowals. Experiment 2 yielded no significant results, likely due to an overpowering of the contextual cues. Experiment 3 rectified this issue and further separated disavowals into the categories 'soften-the-blow' (StB) and 'away-from-insinuated-meaning' (AfIM). The types of denials were found not to be equally effective. Cancellation and disavowal AfIM make the speaker appear to be speaking in better faith than disavowal StB, though the latter is more staining on the target.

Analyzing Commitment Denial from a Multimodal Context

Eda Sükan & Bernardo Rilla

University of Groningen

This presentation examines a case in which the leader of the German right-wing populist party AfD publicly denies a previously incurred commitment. What makes this case distinctive is that the commitment is traced back to a multimodal context. The aim of the presentation is to provide empirical support for commitments undertaken through multimodal means, analyzed from a dialectical perspective. To achieve this, the analysis questions the type of meaning conveyed by pictorial content, specifically whether it should be understood as implicatures or explicatures. The examination of the pictorial material draws on discourse-semantic frameworks, which provide the basis for the pragmatic inferences involved. Finally, the presentation proposes a classification of the type of denial observed in the AfD case.

Denial as a premiss of an accusation of bad faith: a means of defence and an indicatory behaviour

Diane Liberatore

University of Neuchâtel

“First, you pin a derogatory label on me. Then you try to wriggle out of it by claiming it wasn’t derogatory at all. I don’t see how a discussion can move forward with someone acting in such bad faith.” (Case #103, 2020)

Despite the anonymity afforded by online discussion forums such as Forum Sceptique and Madmoizelle, accusing someone of bad faith is a costly act, threatening discursive faces (Goffman 1974, 1996b, 1996a) of the accuser in the eyes of an audience of unrati ed interlocutors who might condemn the aggression that an unjustified accusation would represent (Nozadi et al., 2015). Consequently, the accuser tends to substantiate their standpoint – “You are / this is in bad faith” – by means of premises pointing to behaviours indicative of their interlocutor’s bad faith. Among the premises of the accusation, forum users resort to denial: often by pointing out a denial made by the person accused, which they deem implausible; sometimes by denying having done or said anything themselves, defending themselves against an accusation they deem unfair – itself taken as evidence of the interlocutor’s bad faith. In line with research question 2 of the DARLING project, we will then examine how these instances of denial manifest as arguments in these ordinary discourses of accusations.

Compiling DARLING’s legal corpus: Denials in committee oversight hearings and where to find them

Bruna Paz Schmid

University of Fribourg

Strand 1 of DARLING relies on corpus methods to investigate deniability in natural discourse in terms of its linguistic, rhetorical-argumentative and dialectical features. Strand 1 starts accordingly with the compilation of corpora for annotation in the two chosen genres of legal and political discourse. The talk will discuss the initial phase of data preparation involved in the compilation of DARLING’s legal corpus in terms of the genre of parliamentary corpora and its influence on methods of corpus compilation. In its current form, the legal corpus is comprised of official transcripts from U.S. and Canadian congressional committee oversight hearings. These hearings are investigative in nature and often deal with societally salient issues such as investigations into corruption allegations or, in the U.S. context, nominations to the Supreme Court and other high offices. The likelihood of denials occurring in such a corpus of parliamentary discourse resulting from congressional oversight is thus expected to be high.

Rethinking Plausible Deniability: Implicatures, Threats, and Legal Responsibility

Misha Müller

University of Neuchâtel

Linguistic research following Pinker et al. (2009) has generated extensive conceptual and experimental work on the strategic benefits of implicit communication, notably the speaker’s ability to convey meaning while allowing for plausible deniability.

This paper introduces a fundamental distinction between the denial of truth-conditional content, on which most research focuses, and the denial of threats, which cannot be assessed in such terms and are arguably more committing from the addressee’s perspective. I argue that this distinction has important implications for legal frameworks, where few effective provisions currently exist to hold speakers accountable for implicit communication. This gap stems from the need to preserve the rule of law by minimizing interpretive subjectivity (or “interpretive charity,” as per Lewiński & Oswald 2013: 170), often to the benefit of the accused, even in cases where an implicature is highly plausible in context.

However, when verbal communication enters the domain of threat – as is frequently the case in private, public, and political contexts – different legal provisions come into play, particularly those related to the duty to protect individuals (e.g., Article 28 of the Swiss Civil Code, with equivalents in other jurisdictions). At this critical juncture, a more robust understanding of the pragmatic tools capable of objectifying speaker responsibility in cases of threatening implicatures is not only relevant but urgently needed.

An argumentative approach to deniability: theoretical and experimental considerations

Lou Odermatt

University of Fribourg

Full-fledged denials (see Mazzearella (2023)) are structured in two parts: an objection to meaning attribution (OMA) and an alternative meaning (AM). In current academic debates on deniability, these two parts are often analyzed, in conversation, as being explicit. However, as for communication in general, in day-to-day life, deniability often occurs (at least partly) implicitly. In this presentation, I will theoretically and experimentally justify the relevance of a pragmatic and argumentative approach to deniability, which could provide tools to recognize, reconstruct, and analyze (partly) implicit denials.

Arguments and denials share similarities, such as (i) their internal structure (they are structured in two components), and (ii) their reliance on contextual information (including considerations on ethos) to assess their acceptability. These similarities support the hypothesis that denials have an argumentative function. To gain empirical support for this claim, I conducted two experiments. The results of the first suggest that the acceptability (or plausibility) of a denial is influenced by context and by the acceptability of its AM. The results of the second support the hypothesis of an argumentative function of denials.