Hidden causal genericity in online discourses

Marie Boscaro (Institut Jean Nicod)

October 10, 2025

Contents

Introduction

Introduction : hidden causal genericity on X

(1) We are taught to watch their drinks at bars so we don't get roofied.[1]

(2) Women are blamed for refusing to have sex & blamed for having it.

(3) GEN : P(x) $\rightarrow$ Q(x).

We analyze GENERAL tweets and propose that they display a **hidden genericity causal structure**. (1) and (2) both function **pragmatically as generics** despite that **no standard semantic generic structure.**

[1] All examples in this presentation are drawn from our survey of an opinion-based corpus originally collected by Saha et al. (2019).

Goal of this presentation

Goal

We aim to uncover the **generic causal mechanism** (P $\rightarrow$ Q) at play in GENERAL tweets on X. We provide a formal characterization and distinguish :

- **Cooperative** causal mechanism : it relies on publicly and/or published available information for every social media user (later $SHARED$). Negociated or tacitly grounded. ;
- **Uncooperative** causal mechanism : it relies on the private beliefs, preferences, etc. of social media user (later $PRIVATE$). Non-negociated belief synchronization.

Research questions

RQs

- On what grounds **hidden causal genericity available** whilst not a standard semantic generic construction ?
- What **kind of relation holds between the two properties associated** ? Relies on a fuzzy common ground ?
- In what sense can such **causal constructions be considered uncooperative and/or cooperative** on social media ?

Roadmap

- **Background**: generic constructions & stereotype-based mechanisms
- **Empirical survey**: dataset, methodology, and statistical overview
- **Analysis**: genericity in online discourse - an uncooperative move?
- **Conclusion** & further directions for research

Background : generic & stereotype-based mechanisms

Generic statement : a causal mechanism

Standard analyses of generic statements focus on **salient semantic structures** such as BPs or ISs.

(4) **ISs:** Because if **a girl** has one-night stands, she's a "slut", but if **a boy** does, he's a "lad."

(5) **BPs: Men** fear jail because of the constant threat of rape. **Women** experience that threat everywhere.

Both constructions instantiate a **canonical causal schema**:

(6) GEN: $P(x) \rightarrow Q(x)$ [2]

This raises a key question: *What entitles a speaker to attribute property Q to x once x is known to have property P?* In other words, ***what justifies this causal relationship?***

[2] See, among others, Lawler (1973); Carlson & Pelletier (1996); Cohen (2001); Greenberg (2002); Beyssade et al. (2012); Krifka (2012).

Causality between two properties

Relationships P & Q	Description	Concepts associated
Non-accidental	The relationship between the subject property (P) and the VP property (Q) is **not-accidental**, but derives from P itself (**definition**, **essence**) or from an **abstract rule** linking P to Q.	Analysis "in virtue of" (Greenberg 2002); analytic (Burton-Roberts 1977); essential property (Lawler 1973); definitional (Krifka 2012); nomic; normative reading (Cohen 2000), etc.
Arbitrary	The relationship between P and Q is **contingent**, the result of **empirical generalization**, inductive reasoning based on shared facts or shared situation.	Descriptive reading (Burton-Roberts 1977; Greenberg 2002; Krifka 2012); accidental (Lawler 1973); inductivist (Cohen 2001), etc.

Table: Typology of relationships between P and Q

Causal relationships between properties justified : **normative relationships** or **descriptive/empirical relationships**.

Causality and common ground

Common ground between speakers [3] plays a **crucial role** in establishing and accepting causal relationships between P and Q.

(7) Common ground: the set of propositions presumed to be mutually shared between interlocutors.

According to Krifka (2012), the interpretation of a generic statement—hence the validity of the causal association $P \rightarrow Q$—relies on a model of **common ground** $CG\langle W, I\rangle$, where W denotes the set of worlds compatible with CG (i.e. set of **facts**), and I represents the set of shared **interpretations** (conventions, norms, stereotypical beliefs, etc.).

[3] See Stalnaker (1978, 2002); Farkas & Bruce (2010); Krifka (2015); Geurts (2018); Yalcin (2024).

Interim summary

In the standard semantic framework:

- Generic statements encode a **causal structure** between two properties.
- This causal association relies on a **common ground** : shared **facts** (empirical generalization) or shared **interpretations** (normative/definitional reading).

→ **What happens when causal structure does not rely on this common ground?**

From genericity to stereotypes

(8) **Definition** : statement about a group assuming all person in group share these properties (Schneider 2005).

Stereotypes instantiate similar **causal** schema between a **social kind** (societally recognized group) & an **arbitrary property**, yet the causal link treated as **self-evident** :

(9) **Example:** "**Women** are too emotional for leadership."

(10) Women(x) $\rightarrow$ Too emotional for leadership(x).

From genericity to stereotypes

Stereotypes thus **reproduce the causal structure of** GEN :

- But **does not rely on any common ground** (shared facts, shared norms)
- Sometimes relies on **private or ideological beliefs synchronization requirement between speaker and audience** (**uncooperative**)
- Force a **normative** reading (law-like, non negociated, non factual link) (see Haslanger (2010)) and encode **essencialist beliefs about property causal link** (Haslanger (2010) ; Schneider 2005 ; Anderson et al 2012 ; Cimpian et al 2010).

Furthermore, this process stabilizes **socially dangerous 'essencialisation'**, reinforcing prejudice and belief-bubble effects (Leslie 2007; Gelman & Roberts 2017).

Hidden causal mechanisms

Private beliefs synchronization between conversational participants **implicitly implied by hidden linguistic mechanisms** already explored in recent works in linguistics (see a.o. Camp 2018 ; Dominguez Armas et al 2023 ; McCready & Henderson 2024) with structures such as :

- **Insinuations** (Camp 2018) or **Provocative insinuations** (Dominguez Armas et al 2023)
- **Dogwhistles** (see a.o. McCready & Henderson 2024)
- **Provocative implicatures** (Dominguez Armas et al 2023)

Hidden causal genericity & social media

Our proposal : GENERAL tweets (opinion/factual) encode **hidden genericity causal structure** ($P \rightarrow Q$) that :

- **Does not always rely on a common ground** (conventions, facts, etc)
- Can also operate through **private beliefs synchronization between participants** (like insinuations)
- **Seconde type particularly effective in asynchronous online contexts where it remains unchallenged** leading to strong groupthink effect.

Empirical survey on Twitter (X)

Dataset

Our analysis is based on an **opinion-oriented** corpus consisting of 1378 tweets that were collected originally by Saha et al (2019) through the hastag #YesAllwomen and related keywords (women, men,...). [4]
We performed a double annotation : for **speech act categories** (refining Saha et al (2019)'s schema) and for a **GENERAL statements on Twitter**.

[4] Tulika Saha, S. Saha, P. Bhattacharyya, 2019 'Tweet Act Classification : A Deep Learning based Classifier for Recognizing Speech Acts in Twitter'

First annotation : Speech act annotation schema

We adapted the speech act taxonomy used by Saha et al. (2019), aligning it with the framework proposed in Laurenti et al. (2022 a,b).

Speech acts	**Definition**
Assertions	New raw piece of information p offered to be added to the common ground (CG).
Subjectives	Expressions of opinions, beliefs, preferences and evaluations (partly overlaps with exclamatives). Offer to add p and evaluation on p to CG.
Jussives	Order, wishes, leading to action.
Interrogatives	Information seeking questions.

Table: Speech act categories in Laurenti et al. 2022a,b

Speech acts : examples

We focus here on the two most represented categories in our corpus : **Subjectives** and **Assertives**.

(11) **Subjectives** : Women shouldn't have to be afraid to walk home alone, when it's dark. [5]

(12) **Assertives** : Girls can't turn down an invite to prom without being stabbed to death.

[5] For reminder : all examples analyzed in our presentation are extracted from Saha et al (2019)'s corpus

Background : annotation of generic statements

A number of corpora have previously included annotations for genericity or generality (ACE, GNOME, etc.), yet with **heterogeneous annotation levels and definitions**:

Reference	Annotation-level	Categories
Govindarajan et al. (2019)	**NP-level & S-level**	Episodic / Habitual / Generic
	NP-level	Number, person, POS, determinent,...
	S-level	Dependence relationships, tense, etc.
Methew & Katz (2009)	**S-level**	Generic (=habitual) / Non-generic
Rehbein (2025)	**Clause-level**	Events / States / Generics / Generalizing / Reports / Questions
Friedrich (2015)	**NP-level & Clause-level**	GEN NP vs Non-GEN NP ; Gen Clause vs Non-Gen Clause
Walker et al. (2006)	**S-level**	GEN (generic) / SPC (specific non-generic) / NG (neg. quantifiers) / USP (underspecified)
Louis & Nenkova (2011)	**S-level**	**GENERAL (broad statements) / SPECIFIC (detailed information)**
Herbelot & Copestake (2010)	**NP-level**	Quantifiers: ONE / SOME / MOST / ALL / QUANT

Table: Overview od generic annotation schemes in linguistics and NLP

Generic Annotation Process in Social Media

New: We integrate **GENERAL** statements within an **evidential framework**, and we offer a finer-grained characterization of **how general statements are realized on social media**.
Our annotation proceeds in **two steps** : **identifying** GENERAL statements in our corpus & **categorizing** them based on their morphosyntactic realization on Twitter.

GENERAL annotation : first-level

We treat all GENERAL statements as **evidential constructions** whose source type is **general knowledge**—that is, content presented as arising from **shared, widely accepted, or universal information**, rather than from **direct or mediated testimony**.[6]

Evidence type	Definition
Direct	Information with first-hand evidence (perception, sensorial, experiential)
Relayed	Information with second-hand or third-hand evidence explicitly mentioned by a hyperlink, a hashtag # or mention @
General knowledge	Information presented as grounded in general knowledge - e.g., general truths, cultural stereotypes, community-shared assumptions. Corresponds to a generalized form of testimony, often realized by generic constructions, that does not need any additionnal proof to be accepted as true.
Loose evidence	Information with second-hand or third-hand evidence not mentioned in the tweet
No evidence	When there is no evidence for the information relayed or the tweet is unrelated to the QUD (indicated by #).

Table: Evidence classification

[6] This annotation is part of a broader survey on evidential and source constructions in social media.

GENERAL annotation : second-level

We propose a fine-grained typology of GENERAL statements and we categorize them based on their morpho-syntactic realization on Twitter.

Label	Definitions	Examples
GENERIC	Canonical generic constructions at NP-level (bare plurals, indefinite singulars) or clausal level (habituals, generic statements). Annotation based on standard semantic framework.	Women who espouse feminist ideals and views are afraid to call themselves feminists
QUANTIFICATIONAL	Clauses containing overt quantifiers (all, every), often introduced through hashtags such as YesAllWomen.	#YesAllWomen are still supposed to be flattered by men sexually harassing them
ASSOCIATIVE	Sentences lacking explicit quantifiers or canonical generic marking, but using impersonal we or you pronouns.	Because we are taught to watch our drinks at bars.

Table: Typology of GENERAL constructions (second-level annotation)

Distribution : speech acts

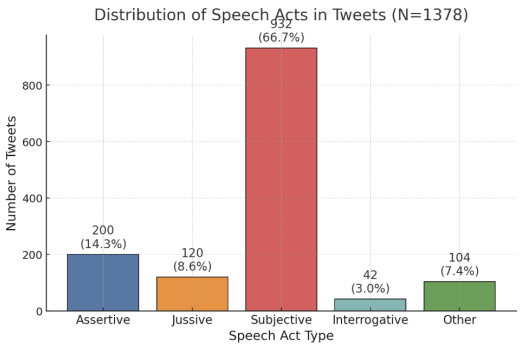


Figure: Speech acts distribution

Majority of annotated tweets are **Subjectives**, as **expected** for an opinion-based corpus.

Distribution : evidence type

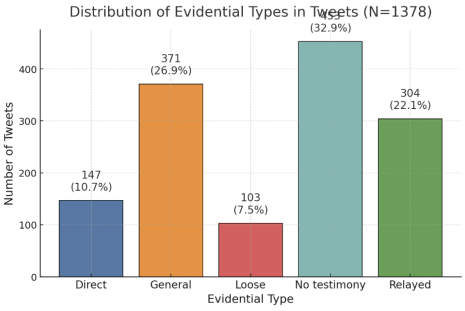


Figure: Evidence type distribution

Most tweets are annotated as lacking evidential anchors, or as drawing on **general-knowledge evidence** (**new and unexpected**).

Distribution : speech acts x evidence type

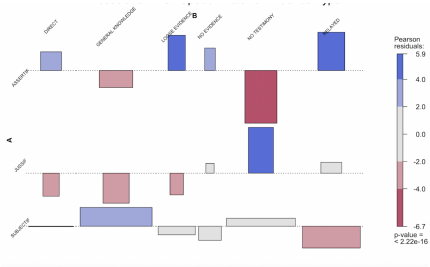


Figure: Association Plot : Speech acts x evidence type

- GK evidence strongly correlated to Subjectives (opinion) & not to Assertives (facts).
- **Opinion based statement rely on GK source** (**unexpected**) - or pretend to do so.

Factual statement rely punctually on GK source among others.

GENERAL types : Distribution

Among GENERAL statements we observe an **uneven distribution** with **GENERIC & ASSOCIATIVE** constructions being predominant.

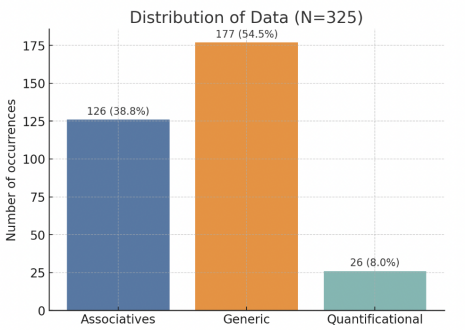


Figure: Distribution of GENERAL constructions in the corpus.

Illustrative examples

(13) Girls are afraid to voice their opinions because "that cute boy won't want me anymore if I'm all feminist." (**Subjective**)

(14) Women (x) $\rightarrow$ Afraid to voice opinion (x).

• This example instantiate an **implicit causal link** between two contingent properties, grounded in private information (beliefs, personal experience, subjective preferences etc) that can be **non-consensual**.

Interim summary

The annotation survey & statistical analysis unveil that :

- GENERAL constructions are **frequent** in **opinion**-based social media discourse.
- Two types of GENERAL constructions mirroring statistical correlations : (i) tied to assertions, convey plain p GK based ; (ii) tied to subjectives (prominent) convey p and associated evaluation GK based
- GENERAL statements manifest in different morpho-syntactic types : **GENERIC**, **QUANTIFICATIONAL** and **ASSOCIATIVE**. Despite this diversity, they all seem to **instantiate an implicit causal mechanism akin to standard generics** which is **justified either by private information or by shared information.**

Analysis : generic statements as uncooperative ?

Features of online discourses

Online discourses possess specific **features** that influence hidden causal dynamics of GENERAL tweets :

- **One-to-many conversation** : speaker posts tweet that is available to an **undeterminate mass audience** $\rightarrow$ (**uncertainty of addressee profile**).
- **Fuzzy common ground** : unacquaintance conversational participants, few conversational moves (like, retweet, answer) signaling conversational commitment, etc. $\rightarrow$ **uncertainty in building common ground**.
- **Asynchronicity** : increase uncertainty grounding ; uncertainty audience's profile.

Two general-based constructions

In our analysis we distinguish two GENERAL tweet types conveying the same causal schema ($P : socialkind(x) \rightarrow Q(x)$) but rest on different informational bases :

- **Cooperative one** $\rightarrow$ based on $SHARED$ information (general knowledge implicitly common ground (Stalnaker (2014); uncertain published information stock forming inductive basis available to every social media users.

  (15) Women are humans. $\Rightarrow$ Women (x) $\rightarrow$ human (x).

- **Uncooperative one** $\rightarrow$ based on $PRIVATE$ information (subjective preferences, private beliefs, etc.) of the speaker and **synchronization requirement** with addressee which is implied by hidden causality.

  (16) You women do the dishes. $\Rightarrow$ Women (x) $\rightarrow$ do the dishes (x).

Structure of online conversation

To spell out this distinction, we offer **formalization of online conversations** with situation theory (a.o. Barwise and Perry (1981), Devlin (2006)) and feature structures framework (a.o. Beyssade and Marandin (2007) ; Ginzburg (2015)). Here **utterance situation**.

$$\left\{ \begin{bmatrix} I_{sp} \\ \textit{PUBLISHED INFO} \left[p : \text{'Women are humans.'} \right] \\ \textit{UNPUBLISHED INFO} \begin{bmatrix} [\textit{SHARED INFO} \ll QUD, \mathsf{CG}_{public}, CGlocal \gg] \\ [\textit{PRIVATE INFO } P_{sp} \ll \mathsf{B}_{sp}\ ^{7} \gg] \end{bmatrix} \end{bmatrix} \begin{bmatrix} I_{ad} \\ \textit{PUBLISHED INFO} \left[\emptyset \right] \\ \textit{UNPUBLISHED INFO} \begin{bmatrix} [\textit{SHARED INFO} \ll ?QUD, CG_{public}, CG_{local} \ll \emptyset \gg \gg] \\ [\textit{PRIVATE INFO } ?P_{ad} \ll \mathsf{B} \gg] \end{bmatrix} \end{bmatrix} \right\}$$

[7] Following Giannakidou & Mari (2021a,b) we consider $\mathcal{B}$ to be both private beliefs of the speaker, but also his subjective preferences

Structure of online conversations

Feature structure modelization aim to describe online conversations grounding process (see a.o. Farkas and Bruce (2010) ; Ginzburg (2015) ; Beyssade and Marandin (2007) ; Gunlogson (2008)). We model information structure of online interaction through two feature structures :

- I_{sp} (**information state of the speaker**) $\rightarrow$ $PUBLISHED$ (the post) and $UNPUBLISHED$ (presumed $SHARED$ information & $PRIVATE$ information of the speaker).
- I_{ad} (**information state's projection of the addresse**)$\rightarrow$ $PUBLISHED$ (the post), and **projection** of $UNPUBLISHED$ (presumed $SHARED$ information (negociated, for CG_{local} tacitly or accommodated CG_{public}) & $PRIVATE$ information projection of the addressee).

$\Rightarrow$ Information state of the addressee is primarily a **projection of the speaker's epistemic state, not an explicitly negotiated common ground**.

Cooperative causal genericity on social media

Cooperative causal genericity relies on $SHARED$: implicit CG (CG_{public} : general knowledge (20), published information available (19)) and/or CG_{local} : explicit CG, negociated through conversation.

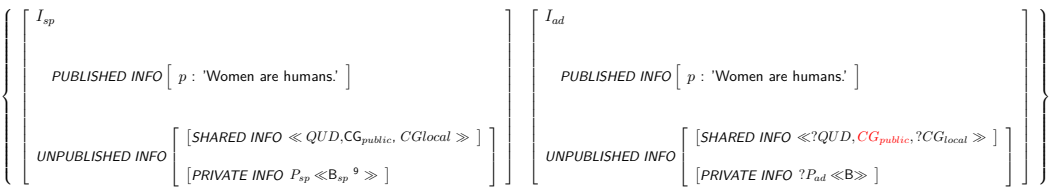
(17) Miss guadeloupe candidates are beautiful. #MissGuadeloupe [8]

(18) Women are humans.

$\rightarrow$ Supports both a **normative** reading (20) and an occasional **descriptive one** (19) : empirical generalization from $PUBLISHED$ informational stock available to social media user (here through hashtag use #MissGuadeloupe).

[8] Extracted from another corpus crise écologique française Kozlowski et al (2021)

Cooperative causal genericity on Twitter



10

[9] Following Giannakidou & Mari (2021a,b) we consider $\mathcal{B}$ to be both private beliefs of the speaker, but also his subjective preferences

[10] Interrogation mark signals the uncertainty of the speaker on the projection of the informational state of the addressee

Uncooperative generics on social media

Uncooperative causal genericity relies on $PRIVATE$: speaker's subjective preference, private beliefs ($\mathcal{B}sp$) projected as **synchronized** with a future addressee ($\mathcal{B}ad$). In this structure, the causal relation $P \rightarrow Q$ depends on a **belief alignment which is presupposed rather than warranted**.

(19) Women are emotional.

$\rightarrow$ **Hidden synchronization of private beliefs licenses the causal link.** Support only **normative reading**.

Uncooperative generics on social media

$$\left\{ \begin{bmatrix} I_{sp} \\ \textit{PUBLISHED INFO} \left[p : \text{'Women are emotional.'} \right] \\ \textit{UNPUBLISHED INFO} \begin{bmatrix} [\textit{SHARED INFO} \ll QUD, \mathsf{CG}_{public}, CGlocal \gg] \\ [\textit{PRIVATE INFO } P_{sp} \ll \mathsf{B}_{sp}\ ^{11} \gg] \end{bmatrix} \end{bmatrix} \begin{bmatrix} I_{ad} \\ \textit{PUBLISHED INFO} \left[p : \text{'Women are emotional.'} \right] \\ \textit{UNPUBLISHED INFO} \begin{bmatrix} [\textit{SHARED INFO} \ll ?QUD, ?CG_{public}, ?CG_{local} \gg] \\ [\textit{PRIVATE INFO } P_{ad} \ll \mathsf{B} \gg] \end{bmatrix} \end{bmatrix} \right\}$$

[11] Following Giannakidou & Mari (2021a,b) we consider $\mathcal{B}$ to be both private beliefs of the speaker, but also his subjective preferences

Rhetorical force of ASSOCIATIVE GENERAL structures

This uncooperative causal mechanism, relying on synchronization of beliefs, can be manipulated by associative structures which possess strong pragmatic effects (cf. Zobel 2014) :

- Foster **cohesion** and present as **default an inclusivity** (I+ associates) between speaker and his unknown audience
- Bias to arbitrary association $P \rightarrow Q$ presented as valid and strenghten with this **cohesion implicature**.

$\rightarrow$ Such structures yield strong **rhetorical effects**, framing subjective and/ or ideological content as **self-evident by default shared with audience** — producing a **perlocutionary acceptance without room for negociation process.**

Normative reading vs. descriptive reading online

On social media, due to online features, the **normative (essentialist) reading of generic statements tends to dominate** on a descriptive one:

- Lack of shared experiences and asynchronous interactions : **few inductive shared basis** (only through limited QUD restricted PUBLISHED information set).
- Hereby **arbitrary causal links ($P \rightarrow Q$) are often interpreted as essential traits of the kind (normative reading)** which produces **essentialist construals of social kinds, paralleling the logic of stereotypes** (Haslanger 2010).

Epistemic hazardous causal construction : Groupthink dynamics

Because online interactions are **asynchronous** :

- Hidden causal structures that remain in the **unpublished layer** cannot trigger an immediate "*Hey, wait a minute!*" challenge; thus **tend to remain unchallenged.**

Because online interactions have **indeterminate audience** which is nonetheless often clustered with speaker through shared private beliefs and preferences (**epistemic bubble** or **echo chamber effects** due to algorithm):

- Echo-chamber dynamics invite **uncritical uptake** of these uncooperative causal constructions, reinforcing **groupthink** and **ideological polarization** between in-group and out-group members.

$\Rightarrow$ Online environments thereby foster the persistence of **uncooperative causal generics sustained by privately synchronized beliefs rather than publicly negotiated evidence**.

Conclusion

Conclusion

We offered an analysis of GENERAL tweets as:

- Encoding a **hidden causal mechanism linking two properties** (GEN-like structure).
- Being **cooperative** when grounded in publicly negotiated - or not - knowledge i.e. $SHARED$, or **uncooperative** when based on private, stereotypical beliefs i.e. $PRIVATE$
- Exhibiting epistemically risky causal patterns **that may reinforce groupthink and stereotypical-beliefs epistemic bubble**.
- Depending on a **fuzzy common ground, less constraining online**, where causal associations remain implicit and/or weakly grounded.

We aimed to reveal the **generic architecture of these statements and their potentially uncooperative causal dynamics**.

Conclusion: Further Research

In future work, we aim to:

- Refine and formalize the notion of fuzzy common ground, and its role in online dynamics.
- Identify and analyze additional forms of uncooperative discourse on social media, and their dependence on this fuzzy common ground.
- Extend this investigation to other one-to-many discourse settings (e.g., political communication, public debates, or alternative social media platforms).

Thank you !

Contact : marie.boscaro@ens.psl.eu

Bibliography

• Saha, T., Saha, S., & Bhattacharyya, P. (2019). 'Tweet act classification: A deep learning based classifier for recognizing speech acts in twitter.' In 2019 International Joint Conference on Neural Networks (IJCNN) (pp. 1-8). IEEE.
• Laurenti, E., Bourgon, N., Benamara, F., Mari, A., Moriceau, V., & Courgeon, C. (2022, June). Give me your Intentions, I'll Predict Our Actions: A Two-level Classification of Speech Acts for Crisis Management in Social Media. In 13th Conference on Language Resources and Evaluation (LREC 2022) (pp. 4333-4343).
• Laurenti, E., Bourgon, N., Benamara, F., Mari, A., Moriceau, V., & Courgeon, C. (2022, July). Speech acts and communicative intentions for urgency detection. In 11th Joint Conference on Lexical and Computational Semantics (* SEM 2022). ACL: Association for Computational Linguistics.
• Govindarajan, V., Durme, B. V., & White, A. S. (2019). Decomposing generalization: Models of generic, habitual, and episodic statements. Transactions of the Association for Computational Linguistics, 7, 501-517.
• Thomas Mathew & Graham Katz. 2009. Supervised Categorization of Habitual and Episodic Sentences. In Sixth Midwest Computational Linguistics Colloquium.