

Language Resources and Evaluation Conference – 13th edition

A (Psycho-)Linguistically Motivated Scheme for Annotating and Exploring Emotions in a Genre-Diverse Corpus

A. Etienne, D. Battistelli, G. Lecorvé

Marseille – June 2022

Emotion annotation: a difficult task

- Which emotional categories (anger, fear, embarrassment, etc.) ?
- Variability of linguistic means to express emotions (e.g. happy, "So many presents for me!")

⇒(psycho-)linguistically motivated annotation scheme

1. (Psycho-)linguistic notions to describe emotions in texts
2. Emotion annotation schemes in NLP
3. Annotation scheme
4. Application of the scheme
5. A few interactions

1. (Psycho-)linguistic notions to describe emotions in texts
2. Emotion annotation schemes in NLP
3. Annotation scheme
4. Application of the scheme
5. A few interactions

(Psycho)-linguistic notions to describe emotions

6 notions to describe emotions

- Expression mode (*how?*)
- Experiencer (*who?*)
- Emotional category (*which emotion?*, anger, embarrassment, etc.)
- Type of emotion (*which type of emotion?*, basic or complex)
- Cause of the emotion (*what causes the emotion?*)
- Consequence of the emotion (*what results from the emotion?*)

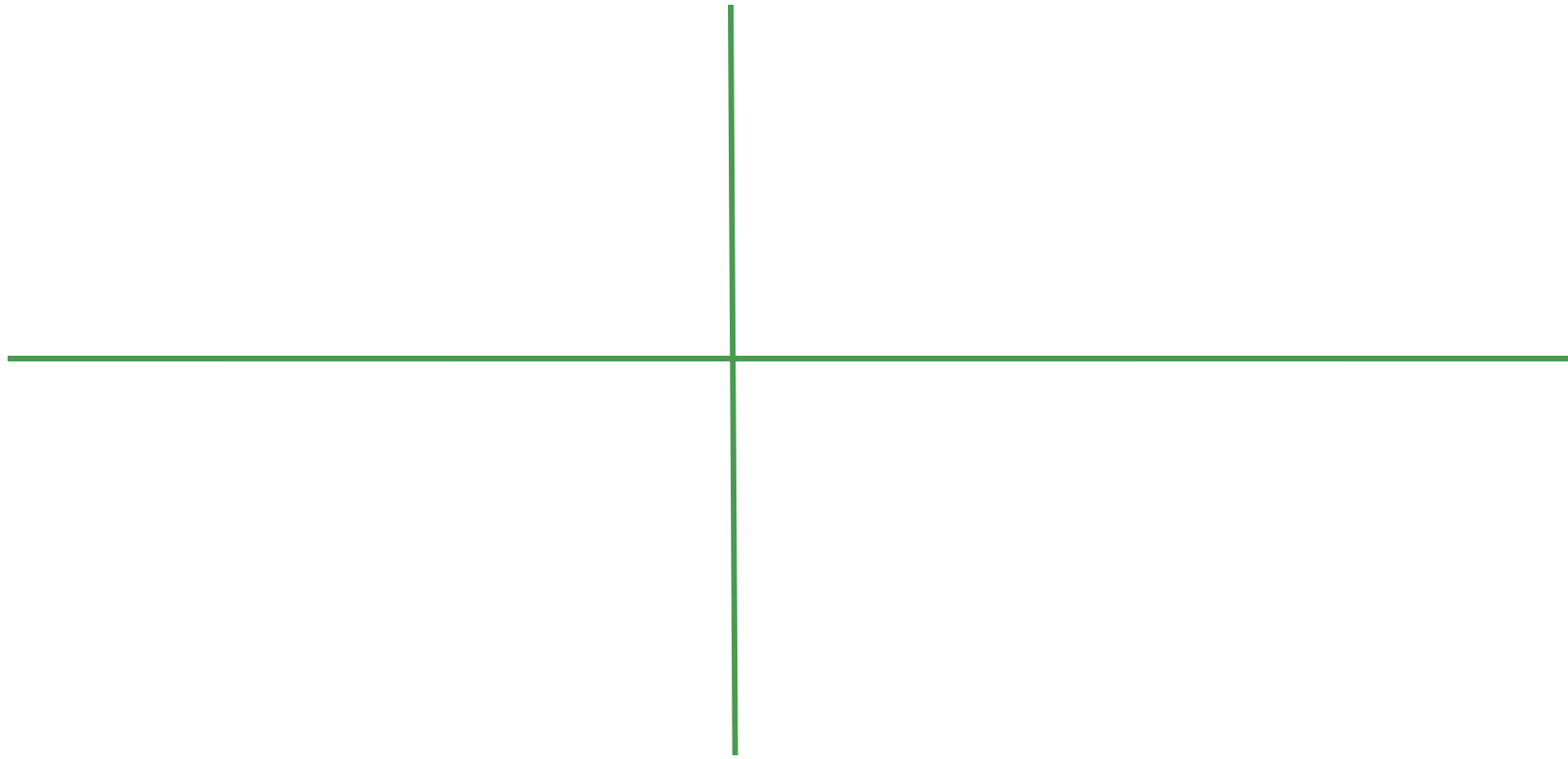
(Psycho)-linguistic notions to describe emotions

Expression modes

- 3 psycho-linguistic expression modes (Creissen and Blanc, 2017)
- 3 linguistic expression modes (Micheli, 2014)

(Psycho)-linguistic notions to describe emotions

Expression modes



(Psycho)-linguistic notions to describe emotions

Expression modes

Labeled emotions

(Micheli, 2014 ; Creissen et Blanc, 2017)

Ex. John is happy.

(Psycho)-linguistic notions to describe emotions

Expression modes

Labeled emotions

(Micheli, 2014 ; Creissen et Blanc, 2017)

Ex. John is happy.

Suggested emotions

(Micheli, 2014 ; Creissen et Blanc, 2017)

Ex. John received a lot of presents.

(Psycho)-linguistic notions to describe emotions

Expression modes

Labeled emotions

(Micheli, 2014 ; Creissen et Blanc, 2017)

Ex. John is happy.

Suggested emotions

(Micheli, 2014 ; Creissen et Blanc, 2017)

Ex. John received a lot of presents.

Behavioral emotions

(Creissen et Blanc, 2017)

Ex. John smiles.

(Psycho)-linguistic notions to describe emotions

Expression modes

Labeled emotions

(Micheli, 2014 ; Creissen et Blanc, 2017)

Ex. John is happy.

Suggested emotions

(Micheli, 2014 ; Creissen et Blanc, 2017)

Ex. John received a lot of presents.

Behavioral emotions

(Creissen et Blanc, 2017)

Ex. John smiles.

Displayed emotions

(Micheli, 2014)

Ex. "Wow! So many presents!" said John.

(Psycho)-linguistic notions to describe emotions

Experiencer

- Linguistic works (Micheli, 2104), (Mathieu, 2005)
- *Who* (person, group of people, etc.) or *what* (animal, etc.) feels the emotion?

Ex. **John** is happy.

(Psycho)-linguistic notions to describe emotions

Emotional categories and emotion types

- From (Blanc and Quenette, 2017) :
 - Basic emotions (anger, disgust, fear, joy, sadness, surprise)
 - Complex emotions (embarrassment, guilt, jealousy, pride + admiration)

(Psycho)-linguistic notions to describe emotions

Cause and consequence of the emotion

- Strong interaction between causality and characters' emotions in text comprehension (Blanc, 2010)

Ex. John is happy. **Yesterday, he discovered a new hiking path.**

Ex. John was feeling depressed. *He did not come to the party.*

1. (Psycho-)linguistic notions to describe emotions in texts
2. Emotion annotation schemes in NLP
3. Annotation scheme
4. Application of the scheme
5. A few interactions

Emotion annotation schemes in NLP

A difficult task: Heterogeneity of emotion annotation schemes

- Two literature reviews:
 - (Bostan and Klinger, 2018)
 - (Öhman, 2020)

Emotion annotation schemes in NLP

Expression modes

- Mention of the diversity of means to express an emotion (e.g. Bostan and Klinger, 2018; Kim and Klinger, 2018; Demszky et al., 2020)
- Observation of linguistic markers of emotions (Alm, 2010)
- No conceptualization, no typology

Emotion annotation schemes in NLP

Expression modes – Type of annotated units

- Text-level annotation (Paroubek et al., 2010; Mohammad, 2012; Mohammad et al., 2018; Fraisse and Paroubek, 2015; Liu et al., 2019; Demszky et al., 2020; Bianchi et al., 2021)
- Sentence-level annotation (Alm et al., 2005; Strapparava and Mihalcea, 2007; Aman and Szpakowicz, 2007; Neviarouskaya and Aono, 2013; Ghazi et al., 2015; Buechel and Hahn, 2017; Öhman et al., 2020)
- Linguistic marker-level annotation (Russo et al., 2011; Neviarouskaya and Aono, 2013; Cheng et al., 2017; Kim and Klinger, 2018; Gui et al., 2018; Schmidt et al., 2021)

Emotion annotation schemes in NLP

Experiencer

- Not mentioned (Aman and Szpakowicz, 2007; Bianchi et al., 2021; Gui et al., 2018; Russo et al., 2011; Ghazi et al., 2015; Liu et al., 2019; Neviarouskaya and Aono, 2013)
- Focus on writer's emotions, without experiencer labeling (Buechel and Hahn, 2017; Demszky et al., 2020; Fraisse and Paroubek, 2015; Mohammad, 2012; Mohammad et al., 2018; Öhman et al., 2020; Paroubek et al., 2010)
- Focus on character's emotions, with experiencer labeling (Cheng et al., 2017; Kim and Klinger, 2018; Schmidt et al., 2021)

Emotion annotation schemes in NLP

Emotional categories and emotion types

- Present in all emotion annotation schemes
- Mostly basic emotions from (Ekman, 1992) or (Plutchik, 1980), cf. (Öhman, 2020)

Emotion annotation schemes in NLP

Cause and consequence of emotion

- Cause of emotion already studied (e.g. Russo et al., 2011; Neviarouskaya and Aono, 2013; Cheng et al., 2017; Kim and Klinger, 2018; Gui et al., 2018; Schmidt et al., 2021)
- No work on consequence of emotion

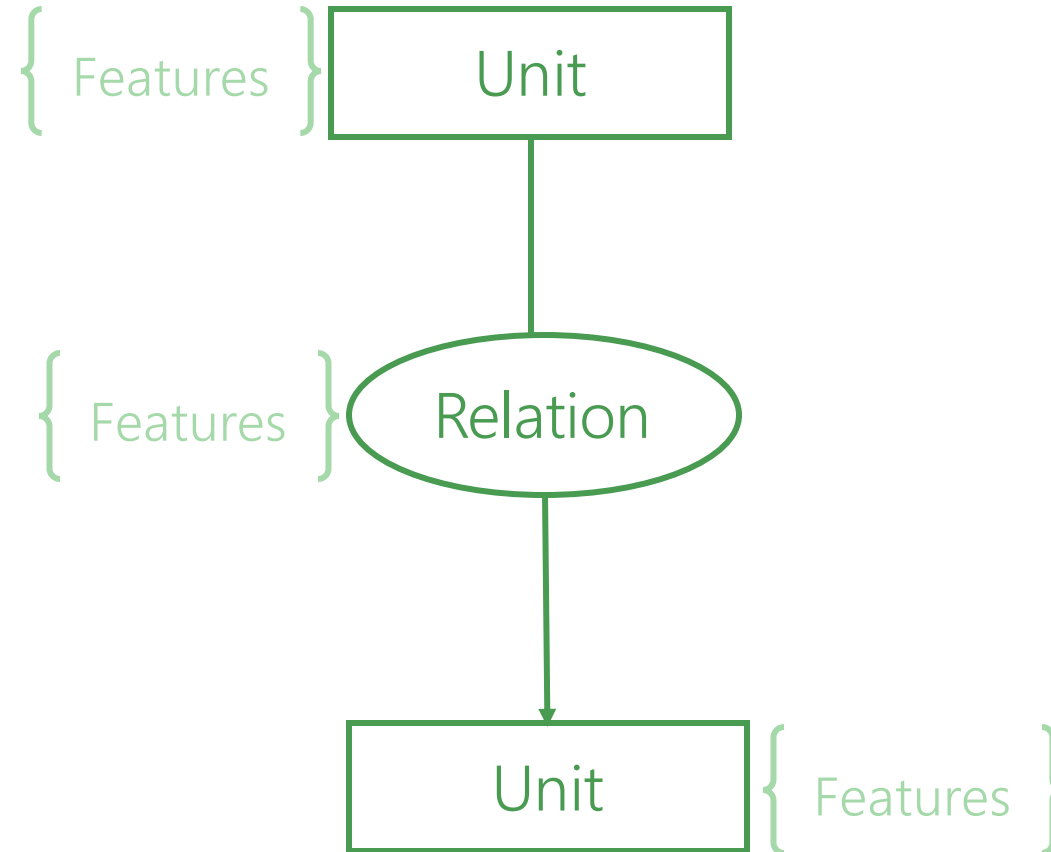
Emotion annotation schemes in NLP

A (psycho-)linguistically motivated scheme

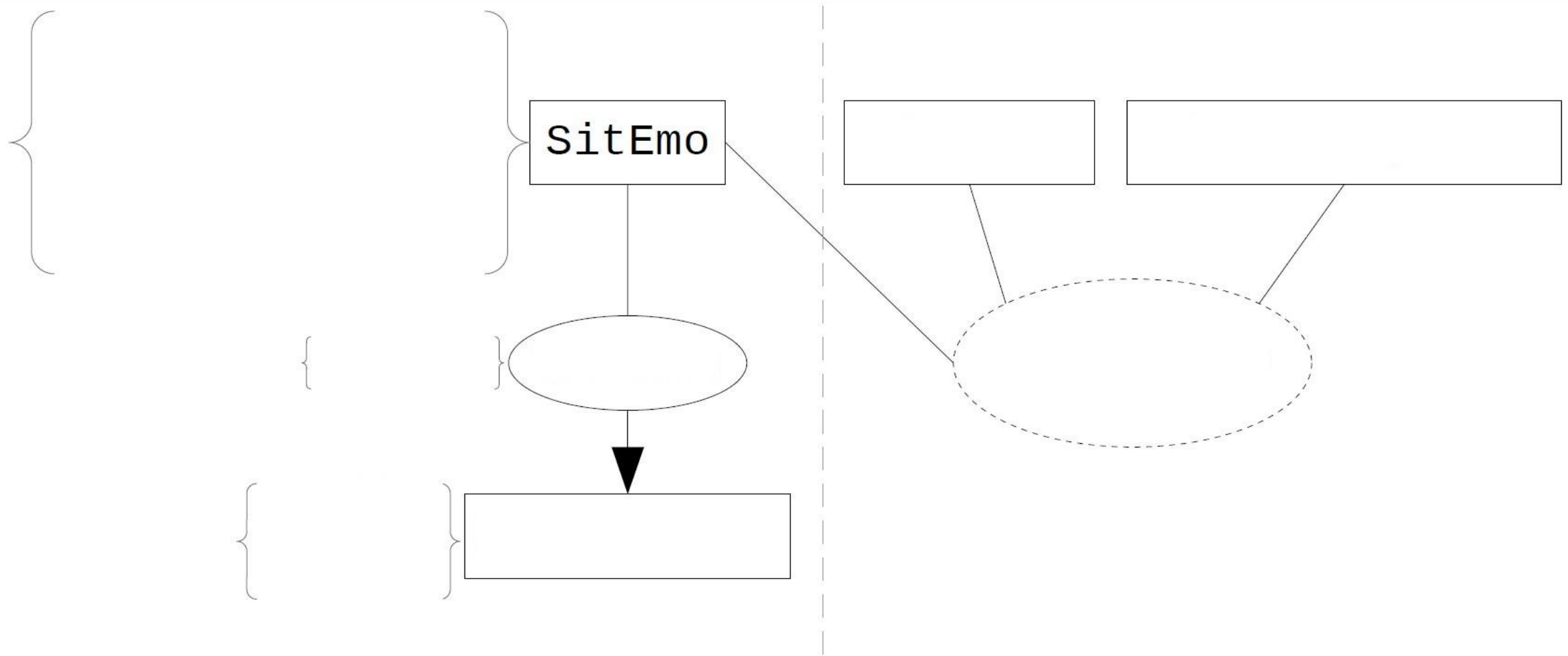
- A scheme that relies on 6 notions to describe emotions:
 - Expression mode
 - Emotional categories
 - Types of emotions
 - Experiencer
 - Cause of emotion
 - Consequence of emotion

1. (Psycho-)linguistic notions to describe emotions in texts
2. Emotion annotation schemes in NLP
3. **Annotation scheme**
4. Application of the scheme
5. A few interactions

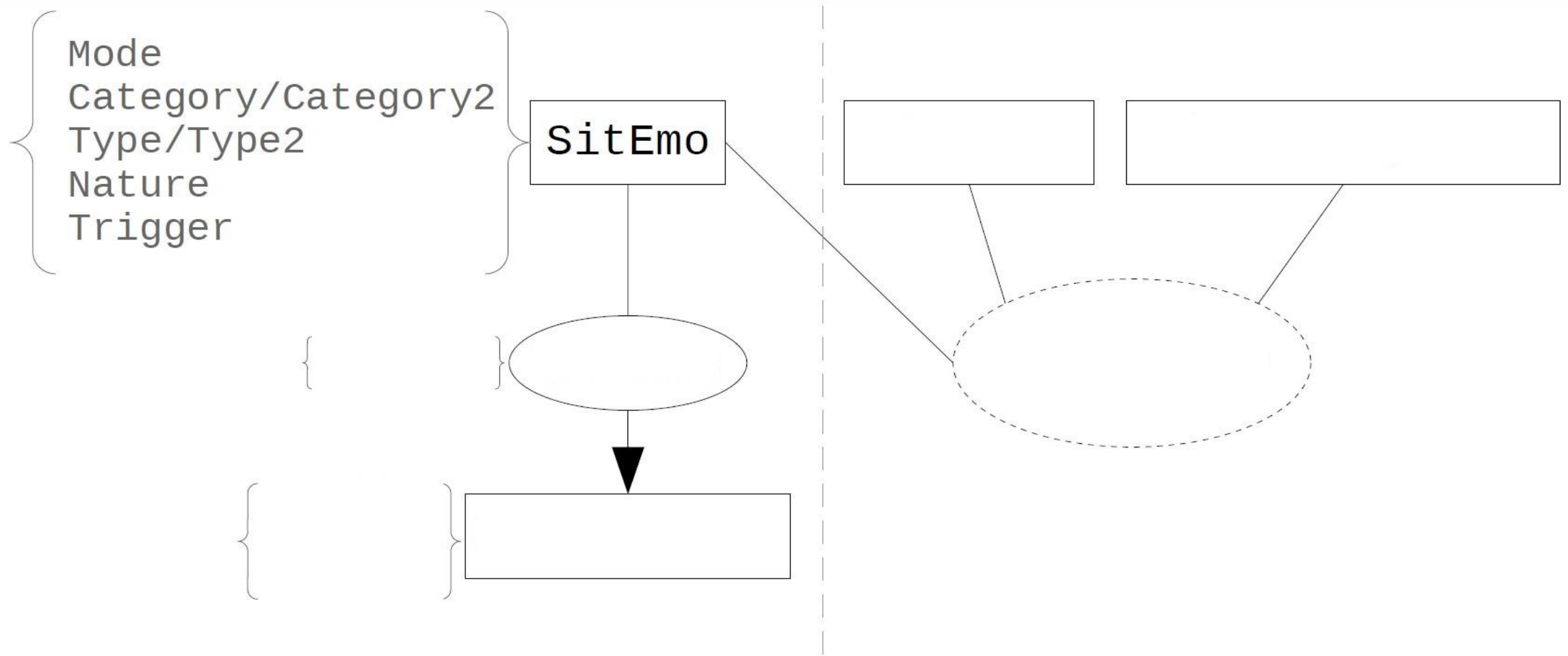
Annotation scheme



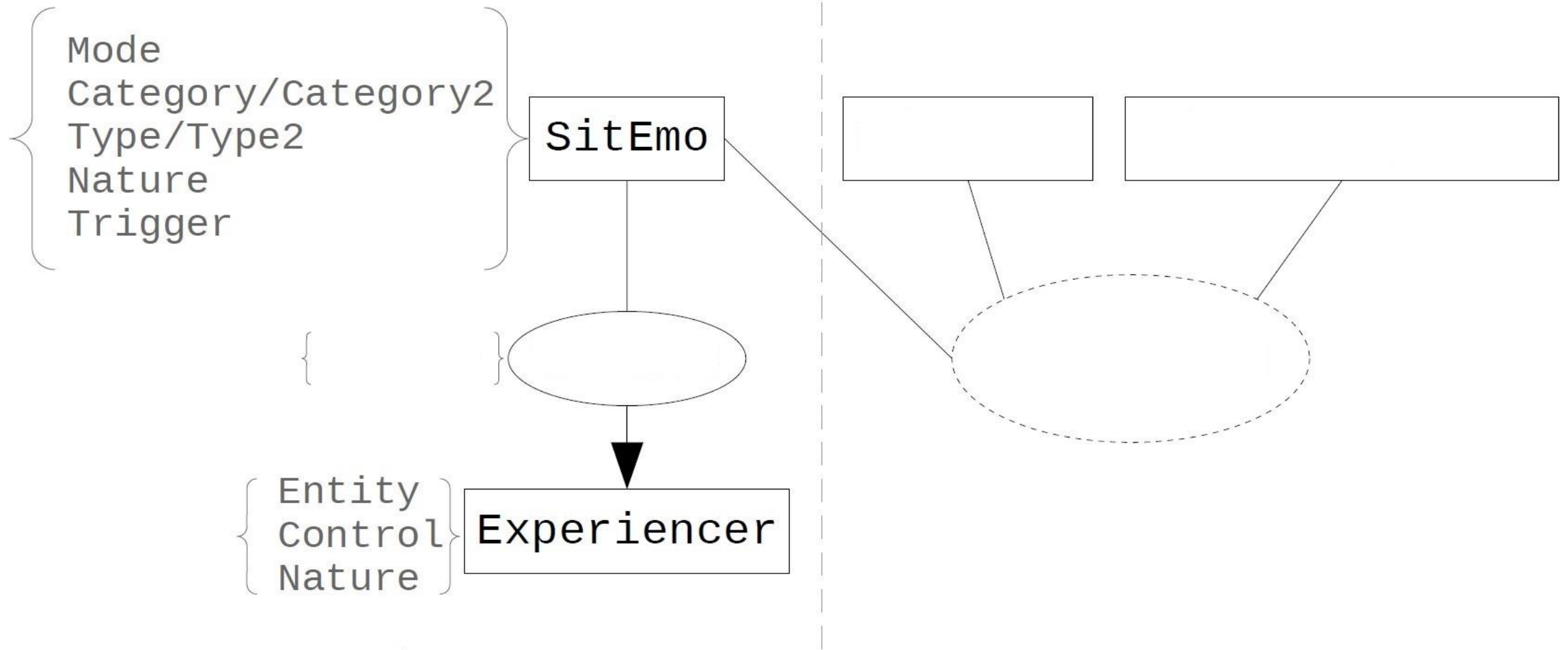
Annotation scheme



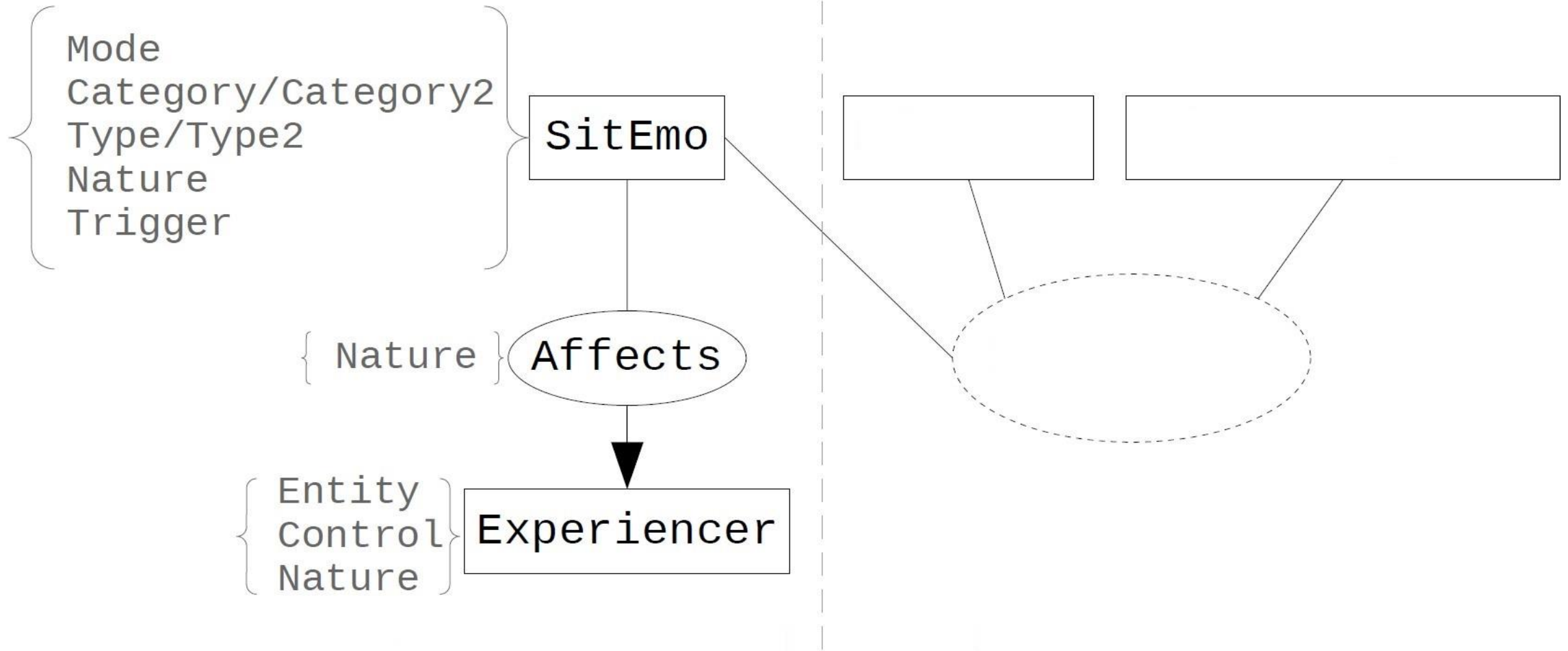
Annotation scheme



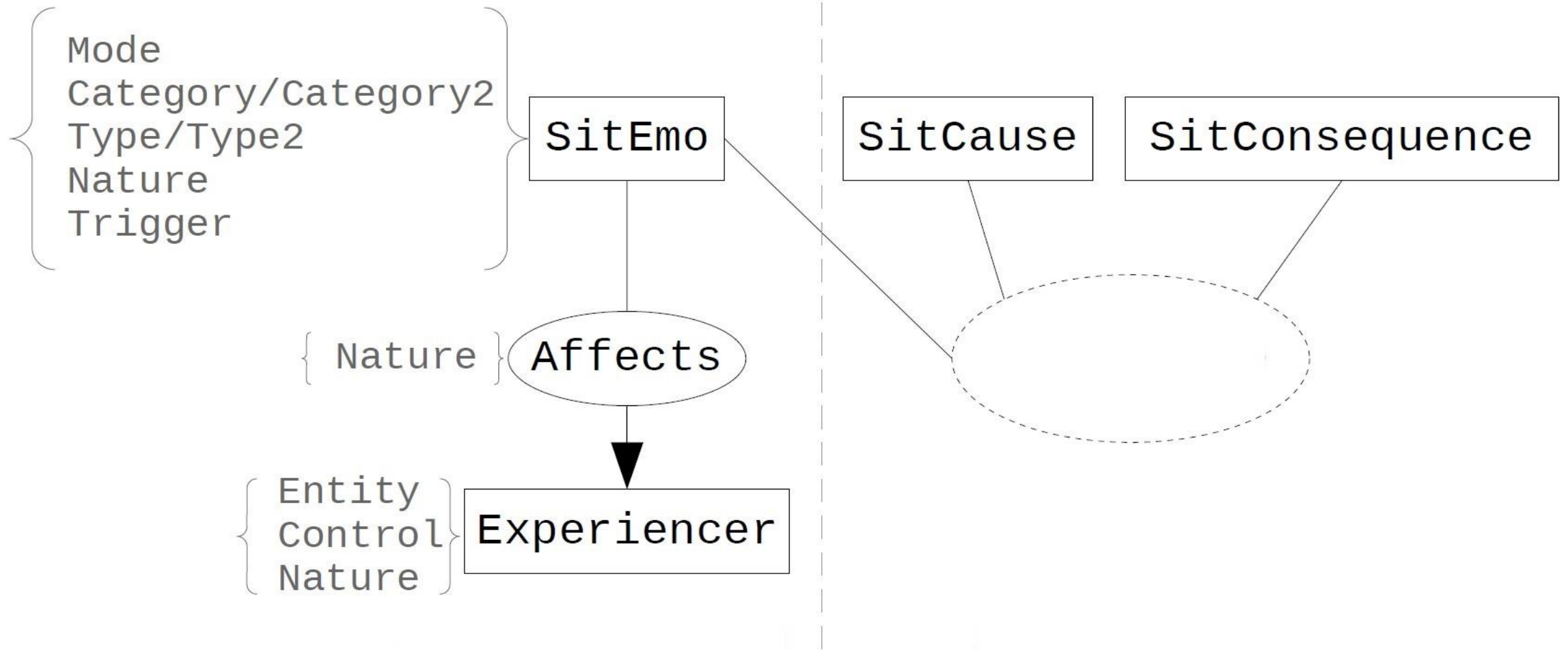
Annotation scheme



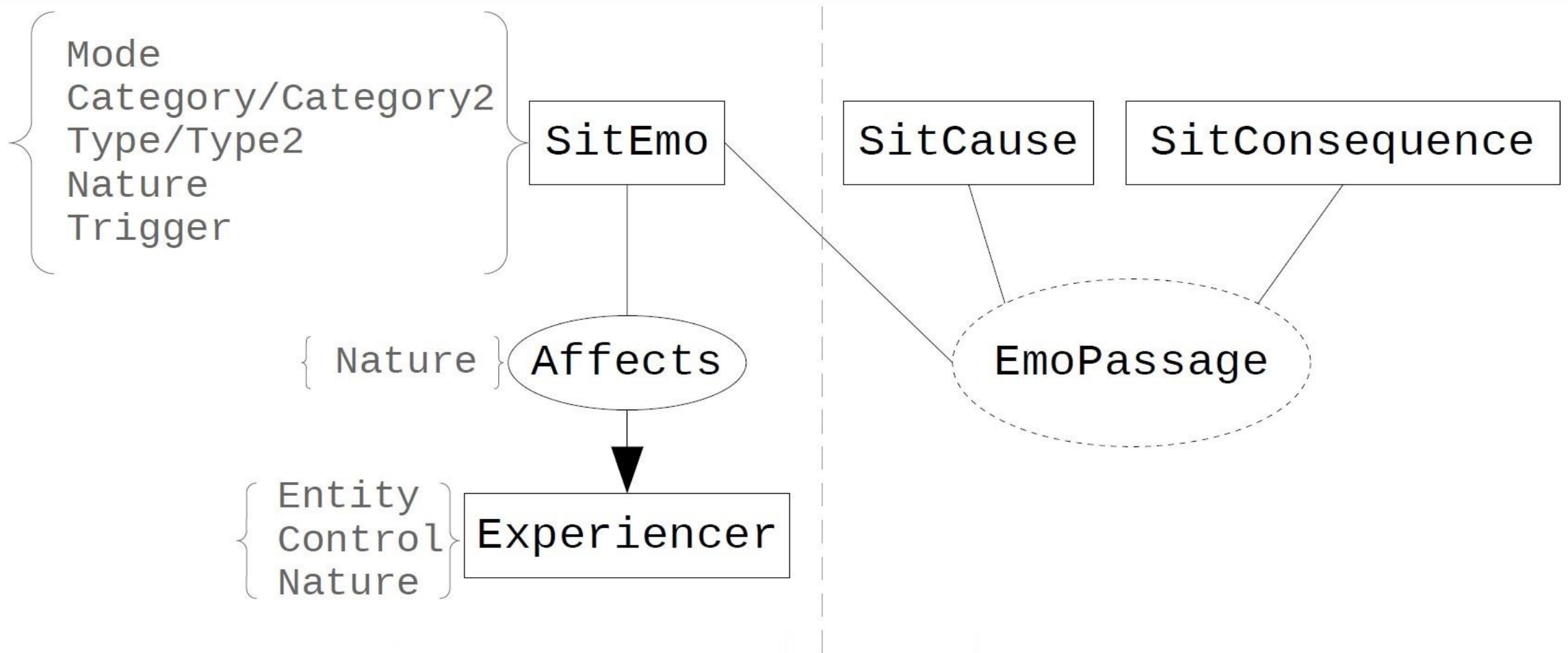
Annotation scheme



Annotation scheme



Annotation scheme



Annotation scheme

Au marché, le paysan les vendit facilement, tandis que le diable ne récolta que moqueries pour ses feuilles fanées. Il en fut tellement vexé qu'on ne le revit jamais plus au pays.

At the market, the farmer sold them[the turnips] easily, while the devil was only mocked for his wilted leaves. He was so offended that he was never seen again in the area.

Example taken from *Le Diable et le paysan*, F. Rabelais.

Annotation scheme

Au marché, le paysan les vendit facilement, tandis que le diable ne récolta que moqueries pour ses feuilles fanées. Il en fut tellement vexé qu'on ne le revit jamais plus au pays.

At the market, the farmer sold them[the turnips] easily, while the devil was only mocked for his wilted leaves. He was so offended that he was never seen again in the area.

Example taken from *Le Diable et le paysan*, F. Rabelais.

Annotation scheme

Au marché, le paysan les vendit facilement, tandis que **le diable** ne récolta que moqueries pour ses feuilles fanées. **Il** en fut tellement vexé qu'on ne le revit jamais plus au pays.

At the market, the farmer sold them[the turnips] easily, while **the devil** was only mocked for his wilted leaves. **He** was so offended that he was never seen again in the area.

Example taken from *Le Diable et le paysan*, F. Rabelais.

Annotation scheme

Au marché, le paysan les vendit facilement, tandis que **le diable** ne récolta que moqueries pour ses feuilles fanées. **Il** en fut tellement vexé qu'on ne le revit jamais plus au pays.

At the market, the farmer sold them[the turnips] easily, while **the devil** was only mocked for his wilted leaves. **He** was so offended that he was never seen again in the area.

Example taken from *Le Diable et le paysan*, F. Rabelais.

Annotation scheme

Au marché, le paysan les vendit facilement, tandis que **le diable** ne récolta que moqueries pour ses feuilles fanées. **Il** en fut tellement vexé qu'on ne le revit jamais plus au pays.

At the market, the farmer sold them[the turnips] easily, while **the devil** was only mocked for his wilted leaves. **He** was so offended that he was never seen again in the area.

Example taken from *Le Diable et le paysan*, F. Rabelais.

1. (Psycho-)linguistic notions to describe emotions in texts
2. Emotion annotation schemes in NLP
3. Annotation scheme
4. **Application of the scheme**
5. A few interactions

Application of the scheme

Annotation procedure

- Glozz platform (Widlöcher and Mathet, 2012)
- 6 expert annotators
- No inter-annotator agreement
- Annotation guide (Etienne and Battistelli, 2021)

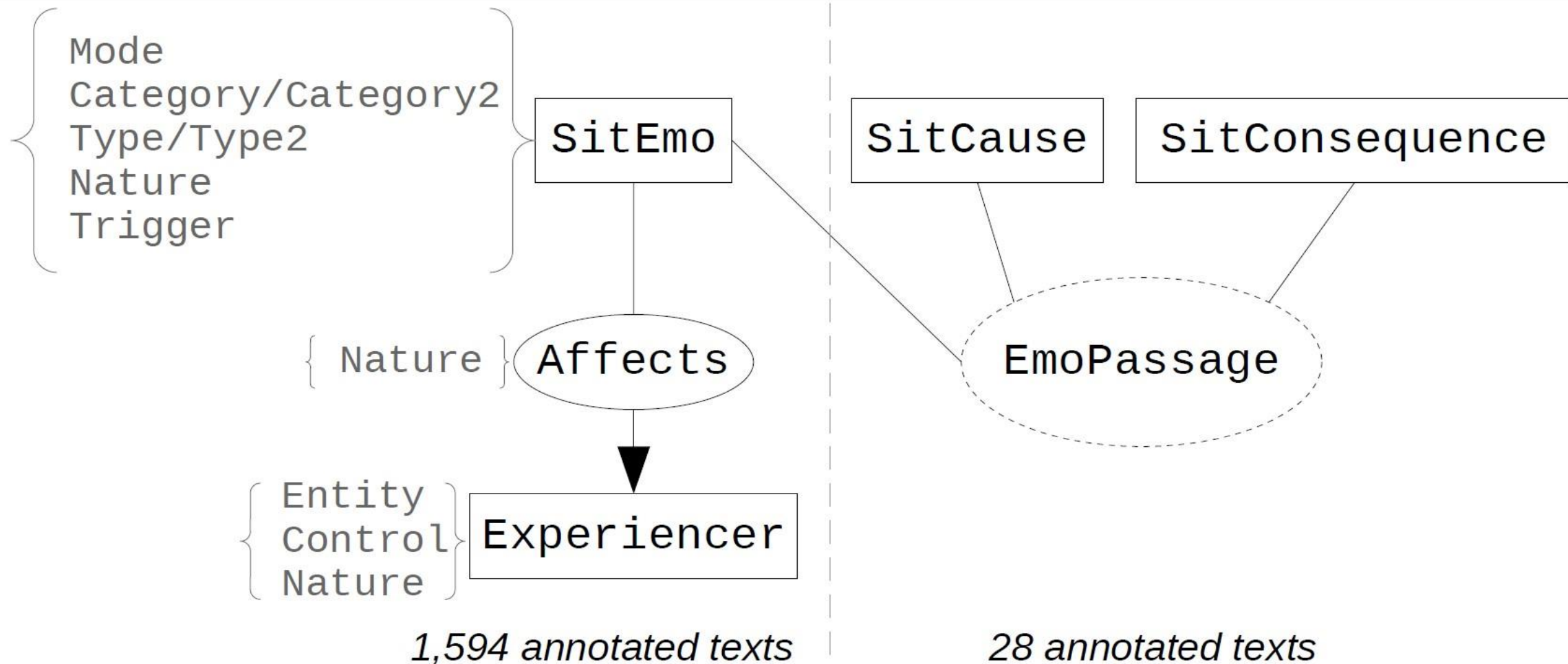
Application of the scheme

Corpus characteristics

Genre	Nb texts	Nb tokens	Mean nb tokens/text
Newspaper	1,566	482K	308
Fiction	17	29K	1,707
Encyclopedia	11	4K	381
Total	1,594	515K	323

- Texts written in French
- Dedicated to 6-14 years old

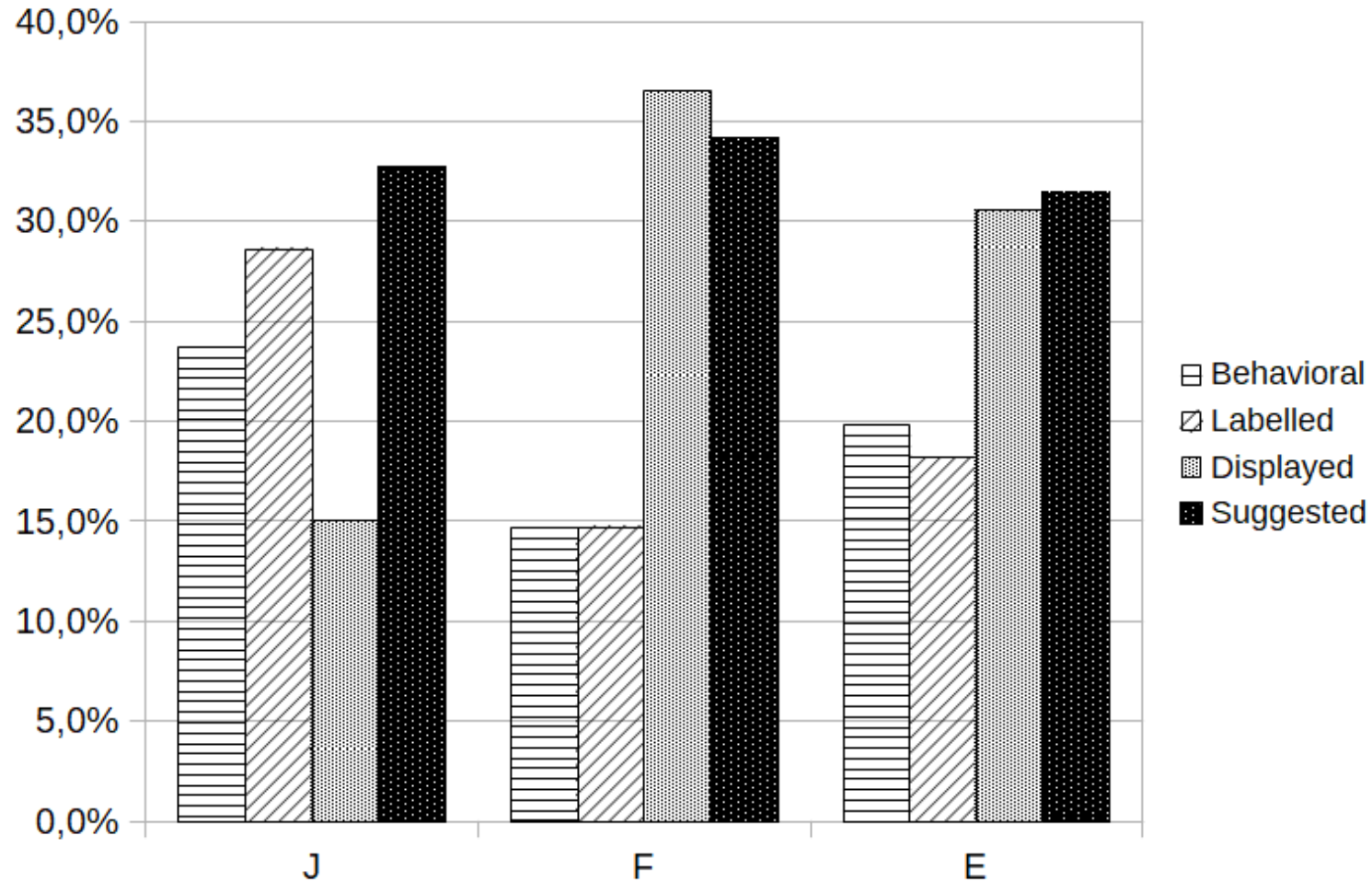
Application of the scheme



1. (Psycho-)linguistic notions to describe emotions in texts
2. Emotion annotation schemes in NLP
3. Annotation scheme
4. Application of the scheme
5. A few interactions

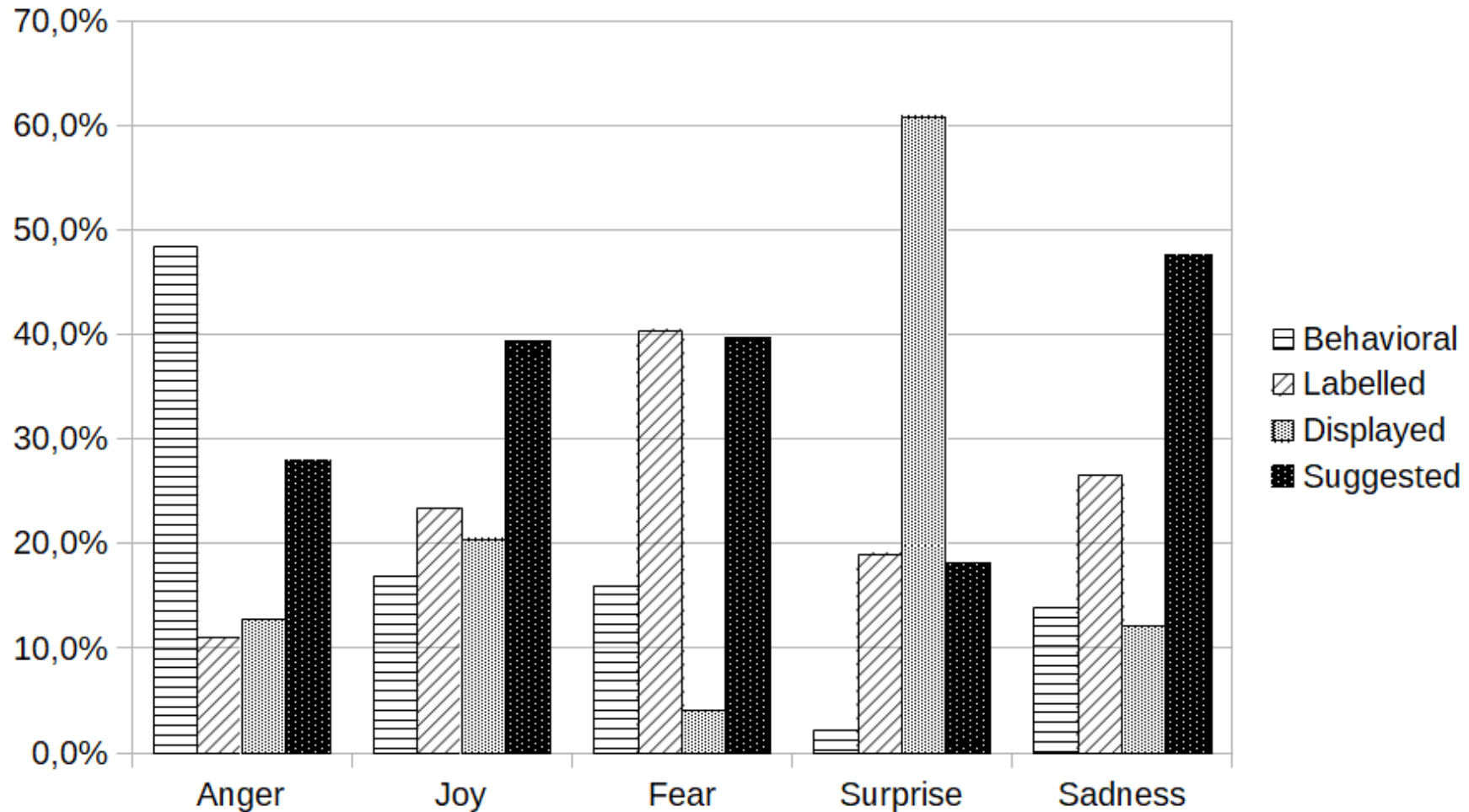
A few interactions

Expression mode and genre of text



A few interactions

Expression mode and emotional category



Conclusion

A scheme...

- ... based on 6 (psycho-)linguistic notions to describe emotion expression in texts
- ... suitable to annotate real textual data of several genres
- ... to explore emotion expression in a systematical and linguistically detailed way

Thank you!

References

- Alm, C. O., Roth, D., and Sproat, R. (2005). Emotions from text: Machine learning for text- based emotion prediction. In Proceedings of Human Language Technology Conference and Conference on Empirical Methods in Natural Language Processing, pages 579–586, Vancouver, British Columbia, Canada, October. Association for Computational Linguistics.
- Alm, C. O. (2010). Characteristics of high agreement affect annotation in text. In Proceedings of the fourth linguistic annotation workshop, pages 118–122.
- Aman, S. and Szpakowicz, S. (2007). Identifying expressions of emotion in text. In International Conference on Text, Speech and Dialogue, pages 196–205. Springer.
- Bianchi, F., Nozza, D., and Hovy, D. (2021). Feel-it: Emotion and sentiment classification for the italian language. In Proceedings of the Eleventh Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis, pages 76–83.
- Blanc, N. (2010). La compréhension des contes entre 5 et 7 ans: Quelle représentation des informations émotionnelles? *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale*, 64(4):256.
- Blanc, N. and Quenette, G. (2017). La production d'inférences émotionnelles entre 8 et 10 ans : quelle méthodologie pour quels résultats ? *Enfance*, 4(4):503–511.
- Bostan, L.-A.-M. and Klinger, R. (2018). An analysis of annotated corpora for emotion classification in text. In Proceedings of the 27th International Conference on Computational Linguistics, pages 2104–2119, Santa Fe, New Mexico, USA, August. Association for Computational Linguistics.

References

- Buechel, S. and Hahn, U. (2017). Emobank: Studying the impact of annotation perspective and representation format on dimensional emotion analysis. In Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics: Volume 2, Short Papers, pages 578–585.
- Cheng, X., Chen, Y., Cheng, B., Li, S., and Zhou, G. (2017). An emotion cause corpus for chinese microblogs with multiple-user structures. ACM Transactions on Asian and Low-Resource Language Information Processing (TALLIP), 17(1):1–19.
- Creissen, S. and Blanc, N. (2017). Quelle représentation des différentes facettes de la dimension émotionnelle d’une histoire entre l’âge de 6 et 10ans ? apports d’une étude multimédia. Psychologie Française, 62(3):263–277. Cognition et multimédia : les atouts du numérique en situation d’apprentissage.
- Demszky, D., Movshovitz-Attias, D., Ko, J., Cowen, A. S., Nemade, G., and Ravi, S. (2020). Goemotions: A dataset of fine-grained emotions. CoRR, abs/2005.00547.
- Ekman, P. (1992). An argument for basic emotions. Cognition and Emotion, 6(3-4):169–200.
- Etienne, A. and Battistelli, D. (2021). Annotation manuelle des émotions dans des textes écrits avec la plateforme Glozz. Research report, MoDyCo ; Université Paris Nanterre, June.
- Fraisse, A. and Paroubek, P. (2015). Utiliser les interjections pour détecter les émotions. In Actes de la 22e conférence sur le Traitement Automatique des Langues Naturelles. Articles longs, pages 279–290.
- Ghazi, D., Inkpen, D., and Szpakowicz, S. (2015). Detecting emotion stimuli in emotion-bearing sentences. In International Conference on Intelligent Text Processing and Computational Linguistics, pages 152–165. Springer.

References

- Gui, L., Xu, R., Wu, D., Lu, Q., and Zhou, Y. (2018). Event-driven emotion cause extraction with corpus construction. In *Social Media Content Analysis: Natural Language Processing and Beyond*, pages 145–160. World Scientific.
- Kim, E. and Klinger, R. (2018). Who feels what and why? annotation of a literature corpus with semantic roles of emotions. In *Proceedings of the 27th International Conference on Computational Linguistics*, pages 1345–1359.
- Liu, C., Osama, M., and de Andrade, A. (2019). DENS: A dataset for multi-class emotion analysis. *CoRR*, abs/1910.11769.
- Mathieu, Y. Y. (2005). Annotation of emotions and feelings in texts. In *Proceedings of the First International Conference on Affective Computing and Intelligent Interaction, ACII'05*, page 350–357, Berlin, Heidelberg. Springer-Verlag.
- Micheli, R. (2014). *Les émotions dans les discours. Modèle d'analyse, perspectives empiriques*. Champs linguistiques. De Boeck Supérieur, Louvain-la-Neuve.
- Mohammad, S., Bravo-Marquez, F., Salameh, M., and Kiritchenko, S. (2018). SemEval-2018 task 1: Affect in tweets. In *Proceedings of The 12th International Workshop on Semantic Evaluation*, pages 1–17, New Orleans, Louisiana, June. Association for Computational Linguistics.
- Mohammad, S. (2012). #emotional tweets. In **SEM 2012: The First Joint Conference on Lexical and Computational Semantics – Volume 1: Proceedings of the main conference and the shared task, and Volume 2: Proceedings of the Sixth International Workshop on Semantic Evaluation (SemEval 2012)*, pages 246–255, Montréal, Canada, 7-8 June. Association for Computational Linguistics.
- Neviarouskaya, A. and Aono, M. (2013). Extracting causes of emotions from text. In *Proceedings of the Sixth International Joint Conference on Natural Language Processing*, pages 932–936.

References

- Öhman, E., Pàmies, M., Kajava, K., and Tiedemann, J. (2020). XED: A multilingual dataset for sentiment analysis and emotion detection. In Proceedings of the 28th International Conference on Computational Linguistics, pages 6542–6552, Barcelona, Spain (Online), December. International Committee on Computational Linguistics.
- Öhman, E. (2020). Emotion annotation: Re-thinking emotion categorization. CEUR Workshop Proceedings, 2865:134–144.
- Paroubek, P., Pak, A., and Mostefa, D. (2010). Annotations for opinion mining evaluation in the industrial context of the DOXA project. In Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC'10), Valletta, Malta, May. European Language Resources Association (ELRA).
- Plutchik, R. (1980). A general psychoevolutionary theory of emotion. In Theories of emotion, pages 3–33. Elsevier.
- Russo, I., Caselli, T., Rubino, F., Boldrini, E., and Martinez-Barco, P. (2011). EMOCause: An easy-adaptable approach to extract emotion cause contexts. In Proceedings of the 2nd Workshop on Computational Approaches to Subjectivity and Sentiment Analysis (WASSA 2.011), pages 153–160, Portland, Oregon, June. Association for Computational Linguistics.
- Schmidt, T., Dennerlein, K., and Wolff, C. (2021). Towards a corpus of historical german plays with emotion annotations. In 3rd Conference on Language, Data and Knowledge (LDK 2021). Schloss Dagstuhl-Leibniz-Zentrum für Informatik.

References

- Strapparava, C. and Mihalcea, R. (2007). SemEval-2007 task 14: Affective text. In Proceedings of the Fourth International Workshop on Semantic Evaluations (SemEval-2007), pages 70–74, Prague, Czech Republic, June. Association for Computational Linguistics.
- Widlöcher, A. and Mathet, Y. (2012). The glozz platform: A corpus annotation and mining tool. In Proceedings of the 2012 ACM symposium on Document engineering, pages 171–180.