

# Unveiling Emotional Landscapes in Plautus and Terentius Comedies: A Computational Approach for Qualitative Analysis

**Daide Picca**

University of Lausanne / Switzerland  
daide.picca@unil.ch

**Caroline Richard**

EDITTA, Sorbonne Université / France  
caroline.richard@sorbonne-universite.fr

## Abstract

This ongoing study explores emotion recognition in Latin texts, specifically focusing on Latin comedies. Leveraging Natural Language Processing and classical philology insights, the project navigates the challenges of Latin's intricate grammar and nuanced emotional expression. Despite initial challenges with lexicon translation and emotional alignment, the work provides a foundation for a more comprehensive analysis of emotions in Latin literature.

## 1 Introduction

Emotion recognition in text, extensively applied to modern languages, has scarcely targeted classical languages like Latin, despite its rich historical and cultural data [Alswaidan and Menai, 2020, Gately, 2023, Korolova et al., 2019]. Recognizing emotions in Latin texts could illuminate classical literature, historical documents, and the evolution of emotional expression.

Latin's intricate grammar, extensive vocabulary, and ancient emotional nuances present unique challenges, requiring sophisticated NLP techniques and cultural understanding [Buzassyova, 2016, Gruber-Miller and Mulligan, 2022]. The limited availability of large, annotated Latin corpora further complicates traditional machine learning applications [Strapparava and Mihalcea, 2008].

This paper addresses these hurdles and explores emotion recognition in Latin texts. We propose a novel method combining NLP techniques and classical philology, extending emotion recognition techniques to Latin language analysis [Pang and Lee, 2008].

## 2 State-of-the-Art

Remarkable strides have been made in computational linguistics and Latin language analysis, including lexicon development [Passarotti, 2016], Medieval Latin Charters annotation [Passarotti,

2019a], Lemlat enhancements [Passarotti, 2019b], and Index Thomisticus Treebank adaptation [Passarotti, 2019c]. However, emotion recognition in classical languages remains relatively untouched.

Essential contributions include Sprugnoli et al. [2020a]'s work on Latin sentiment lexicons and sentiment analysis in Latin poetry [Sprugnoli et al., 2020b]. Studies on other classical languages, like Greek, also offer valuable insights [Yeruva et al., 2020, Pavlopoulos et al., 2022].

Even with these developments, the complexity of Latin's grammar and emotion portrayal makes this a challenging, yet fertile field. A blend of advanced Natural Language Processing (NLP) techniques and a robust understanding of the language's heritage are key to unlocking this potential, promising more profound insights into emotion recognition in classical languages.

## 3 Methodology

### 3.1 Research Design

The research design for this study commences with a quantitative phase, employing NLP techniques such as tokenization and lemmatization, in conjunction with a lexicon-based approach for emotion recognition. It is worth noting that our discourse analysis deviates from conventional norms by adopting a character-based perspective, facilitating the exploration of play dynamics through emotional trajectories Vandersmissen [2019]. In alignment with this perspective, we segment and index the texts according to speaker metadata, thus facilitating an individualized character analysis.

After the quantitative phase, we integrate a qualitative analysis, studying selected Latin texts to understand language and emotion, and devising an emotion coding scheme based on study principles. We then merge quantitative and qualitative results, comparing computational and manual analyses. These findings address the research questions,

deepening our understanding of the issue. This approach sets the foundation for our larger project: developing an emotion lexicon for Latin studies, reducing modern language bias to ensure authentic emotional data extraction.

### 3.2 Data Collection

The data for this study was collected from the Perseus Digital Library<sup>1</sup>. Our attention centers on investigating the genre of Latin comedy, with a specific emphasis on Plautus and Terentius's works. These plays were designated for observation owing to their applicability and donations to the Latin comedy genre and to their relative completeness (indeed most comic plays have been lost or are known by fragments).

The digitization of these texts was already completed by the Perseus Digital Library [Smith et al., 2000], which has undertaken extensive efforts to digitize and preserve classical texts. The data collection process involved downloading the relevant files from the GitHub repository and processing them using a Python script. This script extracts the text content from the XML files, along with the associated metadata. The extracted data is then saved in a structured format (CSV) for further analysis.

### 3.3 Emotion detection on Latin comedies

This ongoing, exploratory project is centered on the investigation of Latin comedy through the application of emotion recognition theory and technology, a realm that promises significant insight into narrative structures and character developments within the genre. Comedy offers a structured medium to analyze the emotional nuances in speeches and compare them with the genre's inherent traits and expectations. Our analysis is directed towards a selected corpus of 26 extant works by the influential Roman playwrights Plautus and Terentius, utilizing advanced computational tools to discern a spectrum of emotions—specifically *anger*, *anticipation*, *disgust*, *fear*, *joy*, *sadness*, *surprise*, and *trust*—embedded in these texts.

The research methodology unfolds in three stages: (1) the integration and expansion of lexical databases, (2) the construction of a lemmatized lexicon utilizing resources from the National Research Council Canada (NRC) for Latin<sup>2</sup>, and (3)

<sup>1</sup>The dataset is freely available at <https://github.com/PerseusDL>

<sup>2</sup><https://saifmohammad.com/WebPages/NRC-Emotion-Lexicon.htm>

meticulously dissecting and interpreting a Latin corpus, which has been subdivided based on speakers' speeches. We implement lemmatization to integrate and enrich the Latin lexicon, reducing word variations to their fundamental form or lemma. It is noteworthy that the lemma detection rate in both stages is approximately 15%, a figure that doubles when excluding lexicon with no emotional connotations.

After the lemmatization stage, the Latin corpus, comprising XML files each representing a Latin literary work, is extracted and parsed. Post-extraction, the speeches are evaluated for emotional content using the prepared NRCLex instance<sup>3</sup>.

Contrary to the assumption that comedic works primarily harbor positive emotions, our preliminary findings reveal a varied emotional terrain, highlighting the complexity within comedic plays. Correlations are observed between specific emotional expressions and character archetypes. However, inherent limitations in the lexicon applied, such as automatic English to Latin translation, alignment of emotions with lemmas based on contemporary English perspectives, and authenticity of employed lemmas, warrant caution. These aspects may engender potential misinterpretations of ancient emotions [Rosenwein, 2010, Konstan, 2016], along with lexical discrepancies, hindering word recognition in the lemmatized corpus, as exemplified by the non-recognition of the verb *metuo* (to fear).

Despite these limitations, the project provides a roadmap for future exploration in this growing field. Efforts are directed towards refining the lexicon and methodology to enhance the assessment of the emotional spectrum within Latin comedy 4, and to advance the broader objective of creating a Latin-specific emotion lexicon, thus enhancing data authenticity and minimizing modern language biases.

### 3.4 Qualitative evaluation of emotion recognition

The computational analysis has yielded substantial insights regarding the emotional strategies employed by Plautus and Terentius to captivate their audiences. It validates the pronounced prevalence of *joy*, thereby affirming the comedic essence of the genre. Furthermore, a notable correlation exists

<sup>3</sup>GitHub repository available here: <https://github.com/CarolineRichard/ENCODEM.git>

between the emotion displayed and the character archetypes, suggesting that Plautus meticulously crafted character personalities and roles to evoke specific emotional responses from his audience. This endeavor has thus substantially broadened our comprehension of the intricate interplay between emotion and language within Roman New Comedy. Initial examinations reveal that comedies adhere to certain core emotional motifs, with *surprise* emerging as the most predominantly depicted and universally shared sentiment. This finding aligns with the inherent narrative logic of comedies, wherein the plot revolves around unforeseen twists, deceptive maneuvers, and mistaken identities. Emotions such as *anger* and *fear* also figure prominently in the narrative landscape. Certain characters appear to be consistently characterized by these dual emotions, such that one is seldom portrayed without the other. This pattern is discernible in characters like *Simon* from *Andria* or *Antiphon* from *Stichus*. The characters embodying these paired emotions often assume pivotal roles in the narrative, such as the *adulescens* (young man) or the *senex* (old man) as shown in Figure 4 in the Appendix.

The intricate interplay between the dual roles in the drama manifests itself through the core dynamic tension between *fear* and *anger*.

From an emotional perspective, characters can be dichotomized into two groups:

- those who experience a broad emotional spectrum.
- others who are defined by single or dual predominant emotions.

The bifurcation of emotional responses can be attributed to the alignment of specific characters with particular emotional types. For instance, as shown in Figure 3 in the Appendix, a majority of the slaves typically display a limited array of emotions, commonly *fear*, *anger*, or *joy*.

In a similar vein, the *parasitus* character is primarily associated with emotions of *anger* or *fear*, with a scant expression of other emotions. Stereotypical emotions in comedy may hint at social representation trends. A character's social status might correlate with the type and complexity of emotions they express. However, even within this framework, emotional responses exhibit significant variations within the same character archetype. For instance, within the demographic of elderly men, certain characters are solely associated with *fear*

and *anger*, whereas others predominantly display *surprise*. This pattern is accentuated by the character discrepancies between the *pater lenis* (gentle father) and the *pater durus* (harsh father), as observed in the *Heautontimoroumenos* (refer to Lhostis [2019] and Figure 2 in the Appendix for further details).

Across various plays, there is a discernible consistency in the characterization: characters predominantly characterized by *anger* and *fear*; those largely exhibiting *surprise*, and others manifesting a diverse emotional range. Within this last group, *joy* tends to be the most prevalent emotion. For example, the narrative of the *Mostellaria* revolves around two young men primarily associated with *anger* and *fear*, while the characters of *Father Teuropides* and *Philematia* the freed courtesan is dominated by *surprise* (See Figure 1 in the Appendix). Other characters display a blend of emotions.

The emotional distribution among characters does not necessarily correspond to their degree of involvement in the narrative arc. For example, in *Mostellaria*, despite the central role of the character *Trianon*, his emotional display is subdued and not polarized. However, this appears more aligned with a distribution based on the characters' roles within the dramatic schema: characters in conflict tend to display polarised emotions, whereas supporting or ancillary characters exhibit a more varied, non-polarised emotional range. This pattern is discernible in plays such as *Stichus*, *Poenulus* and *Mostellaria*, among others.

Quantitative research underscores recurring patterns in the dramatic construction of comedy and stock characters, specifically in the works of Plautus and Terentius, which deftly employ complex emotion networks to enhance the dynamicity of their plays. Unexpectedly, each play exhibits a unique global emotion network thereby suggesting that each play dynamic is distinct, regardless of their stereotypical characters and plots. This emotion-centric interaction is integral in shaping audience reception. The balance struck between standard emotional archetypes, such as the *pater durus*' *anger*, and an innovative emotional dynamic indicates the nuanced comical effects.

Traditional analysis, which emphasizes plot progression and dramatic dynamics, may overlook these emotional nuances. Therefore, this study advocates an alternate perspective that emphasizes the emotional interaction between characters.

## 4 Future Works

This proposal seeks to build upon our initial analysis of character dialogues, with the objective of developing an enhanced lexicon, rooted in the foundational NRC-Emolex model. This process includes meticulous data sanitization and augmentation of emotional markers. After refining the model, the next step is to study two plays, comparing manual and automatic emotional annotations, to further improve emotion recognition.

The indispensable preliminary discourse analysis provides a foundational understanding, vital to the formulation of a specialized emotion lexicon for Latin textual studies. By doing so, the proposal aims to reinforce the efficiency of emotion detection and bolster the reliability and authenticity of the extracted emotional data.

The complexity of emotional semiotization necessitates that we do not solely depend on specific emotion-related lemmas, given their inherent instability and context-dependence Micheli [2014].

Utilizing emotional markers derived from a Latin corpus, along with phraseological characteristics, will enrich our lexicon via a thematic, rather than strictly lexical approach. This will enable a more precise comprehension of emotions, allowing for an accurate assessment of the emotion network and a nuanced understanding of emotional representation.

The proposed project, thus, marks a significant step towards an exhaustive tool for deep investigation of emotions within Latin literature. This innovative endeavour is set to amplify our understanding of the emotional dimensions present within these foundational texts.

## References

- N Alswaidan and MEB Menai. A survey of state-of-the-art approaches for emotion recognition in text. *Knowledge and Information Systems*, 2020. doi: 10.1007/s10115-020-01449-0. Query date: 2023-06-21 21:34:37.
- Ludmila Buzassyova. The 'phonetic complex' in renaissance latin grammar petrus ramus's dichotomies and their reflections in two vernacular grammatical texts. *Graeco-Latina Brunensia*, 21(2):81–98, January 2016. doi: 10.5817/GLB2016-2-8.
- Jane Gatley. Cultural capital, curriculum policy and teaching latin. *British Educational Research Journal*, 49:174–185, 2023.
- John Gruber-Miller and Bret Mulligan. Latin vocabulary knowledge and the readability of latin texts: A preliminary study. *New England Classical Journal*, 2022. doi: 10.52284/necj.49.1. article.gruber-millerandmulligan. URL <https://dx.doi.org/10.52284/necj.49.1.article.gruber-millerandmulligan>.
- Konstan. Their emotions and ours: A single history? *L'Atelier du Centre de recherches historiques*, 16, 2016. doi: <https://doi.org/10.4000/acrh.6756>. URL <http://journals.openedition.org/acrh/6756>. Online; accessed 27-June-2023.
- Nataliia Korolova, Oksana Koshchii, and Valentyna Myronova. The latin language as a universal cultural code. *Journal of History Culture and Art Research*, 8:278–290, 2019.
- Nathalie Lhostis. Le langage de la sagesse dans l'heautontimoroumenos de térence. *Vita Latina*, 199 (1):171–195, 2019. doi: 10.3406/vita.2019.1909. URL [https://www.persee.fr/doc/vita\\_0042-7306\\_2019\\_num\\_199\\_1\\_1909](https://www.persee.fr/doc/vita_0042-7306_2019_num_199_1_1909).
- Raphaël Micheli. *Les émotions dans les discours: modèle d'analyse, perspectives empiriques*. Champs linguistiques. de Boeck Duculot, Louvain-la-Neuve, 2014. ISBN 978-2-8011-1738-5.
- B. Pang and L. Lee. Opinion mining and sentiment analysis. *Foundations and Trends in Information Retrieval*, 2(1-2), 2008.
- Marco Passarotti. Building a word formation lexicon for latin. In *LREC*, 2016. URL <https://www.aclweb.org/anthology/L16-1681.pdf>.
- Marco Passarotti. Annotating medieval latin charters. In *Proceedings of the NoDaLiDa 2019 Workshop on Processing Historical Language*, 2019a. URL <https://www.aclweb.org/anthology/W19-4718.pdf>.
- Marco Passarotti. Enhancing the latin morphological analyser lemlat with an onomasticon. In *Proceedings of the NoDaLiDa 2019 Workshop on Processing Historical Language*, 2019b. URL <https://www.aclweb.org/anthology/W19-4719.pdf>.
- Marco Passarotti. Converting the index thomisticus treebank into universal dependencies. In *Proceedings of the NoDaLiDa 2019 Workshop on Processing Historical Language*, 2019c. URL <https://www.aclweb.org/anthology/W19-4720.pdf>.
- J Pavlopoulos, A Xenos, and D Picca. Sentiment analysis of homeric text: The 1st book of iliad. *Proceedings of the Thirteenth Language Resources and Evaluation Conference*, 2022. URL <https://aclanthology.org/2022.lrec-1.765/>.
- Barbara Rosenwein. Problems and Methods in the History of Emotions. *Passions in Context I. International Journal for the History and Theory of Emotions*, 1, January 2010.

- David A Smith, Jeffrey A Rydberg-Cox, and Gregory R Crane. The perseus project: A digital library for the humanities. *Literary and Linguistic Computing*, 15 (1):15–25, 2000.
- Rachele Sprugnoli, Marco Passarotti, Daniela Corbetta, and Andrea Peverelli. Creating, evaluating and extending sentiment lexicons for latin. In *Proceedings of The 12th Language Resources and Evaluation Conference*, pages 3078–3086, 2020a.
- Rachele Sprugnoli, Marco Passarotti, Daniela Corbetta, and Andrea Peverelli. Odi et amo: Creating, evaluating and extending sentiment lexicons for latin. In *Proceedings of the 12th Conference on Language Resources and Evaluation (LREC 2020)*, pages 3078–3086. European Language Resources Association (ELRA), 2020b.
- C. Strapparava and R. Mihalcea. Learning to identify emotions in text. In *Proceedings of the 2008 ACM symposium on Applied computing*, 2008.
- Marc Vandersmissen. Discours des personnages féminins chez Seneque. *Collection Latomus*, 2019. URL [https://www.academia.edu/38169453/Discours\\_des\\_personnages\\_f%C3%A9minins\\_chez\\_S%C3%A9n%C3%A8que\\_Approches\\_logom%C3%A9triques\\_et\\_contrastives\\_dun\\_corpus\\_th%C3%A9%C3%A2tral](https://www.academia.edu/38169453/Discours_des_personnages_f%C3%A9minins_chez_S%C3%A9n%C3%A8que_Approches_logom%C3%A9triques_et_contrastives_dun_corpus_th%C3%A9%C3%A2tral). ISBN: 9789042937970.
- Vijaya Kumari Yeruva, Mayanka Chandrashekar, Yungyung Lee, Jeff Rydberg-Cox, Virginia Blanton, and Nathan A Oylar. Interpretation of sentiment analysis in aeschylus’s greek tragedy. In *Proceedings of LaTeCH-CLfL 2020*, pages 138–146, 2020.

## **A Appendix: Figures**

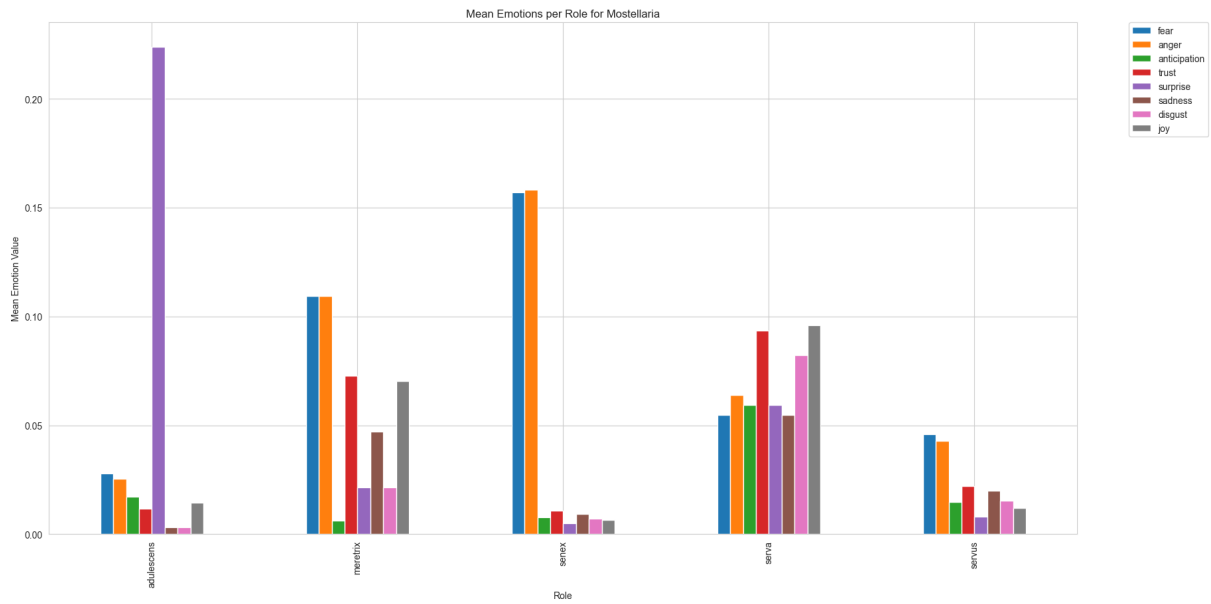


Figure 1: Emotional Distribution in Mostellaria

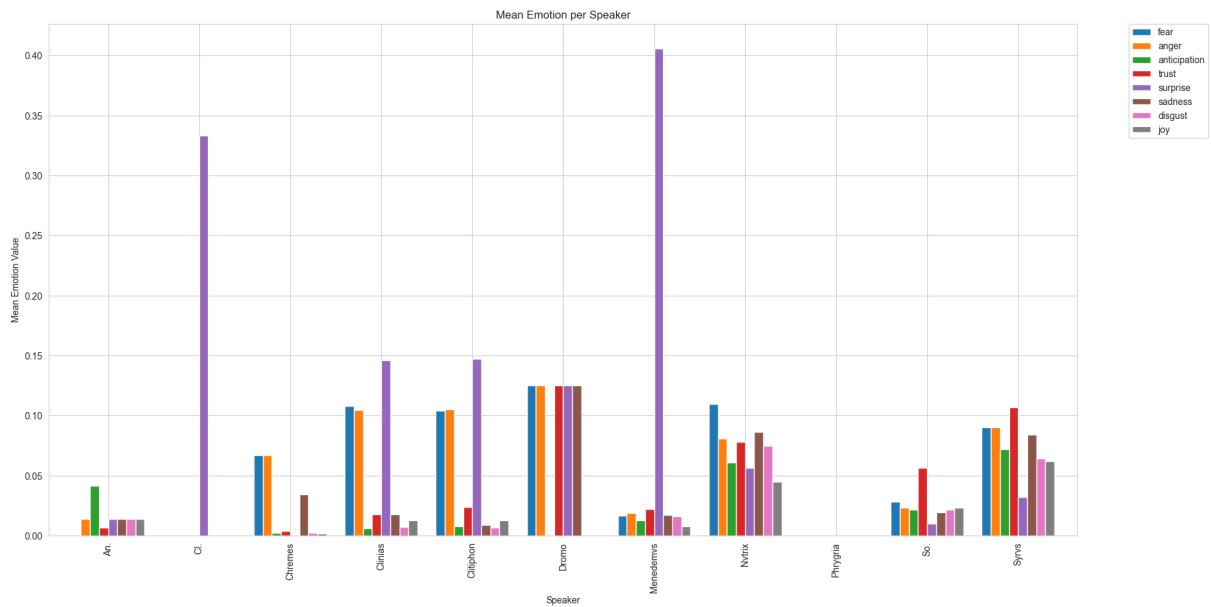


Figure 2: Emotional Landscape in the Heautontimoroumenos (Terentius)

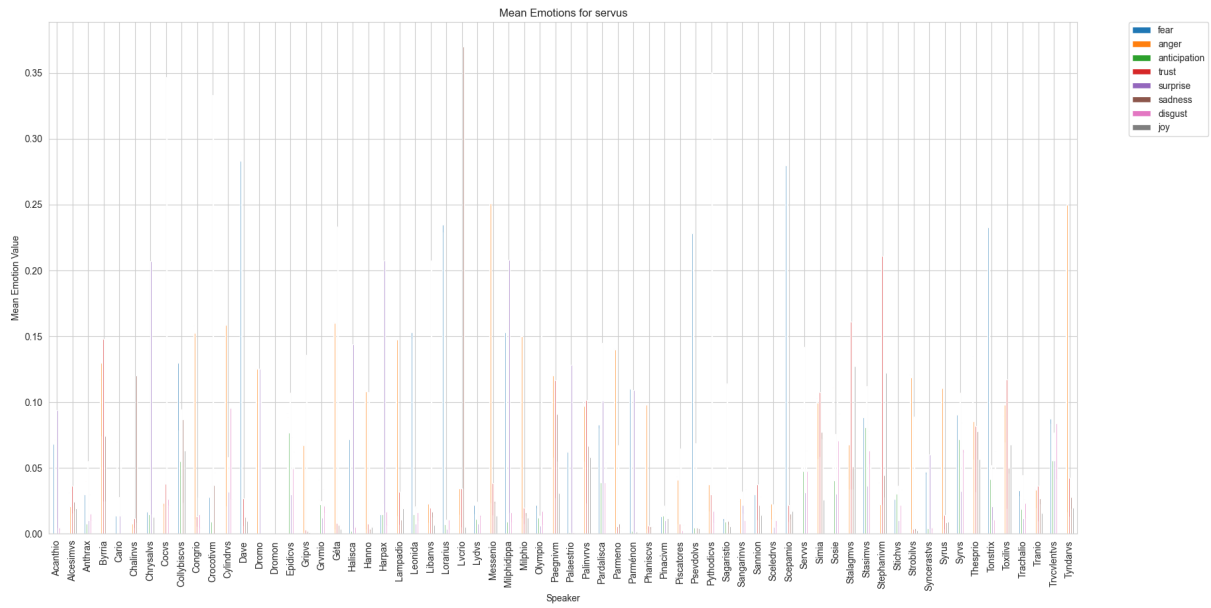


Figure 3: Representation of slaves' Emotional Spectrum

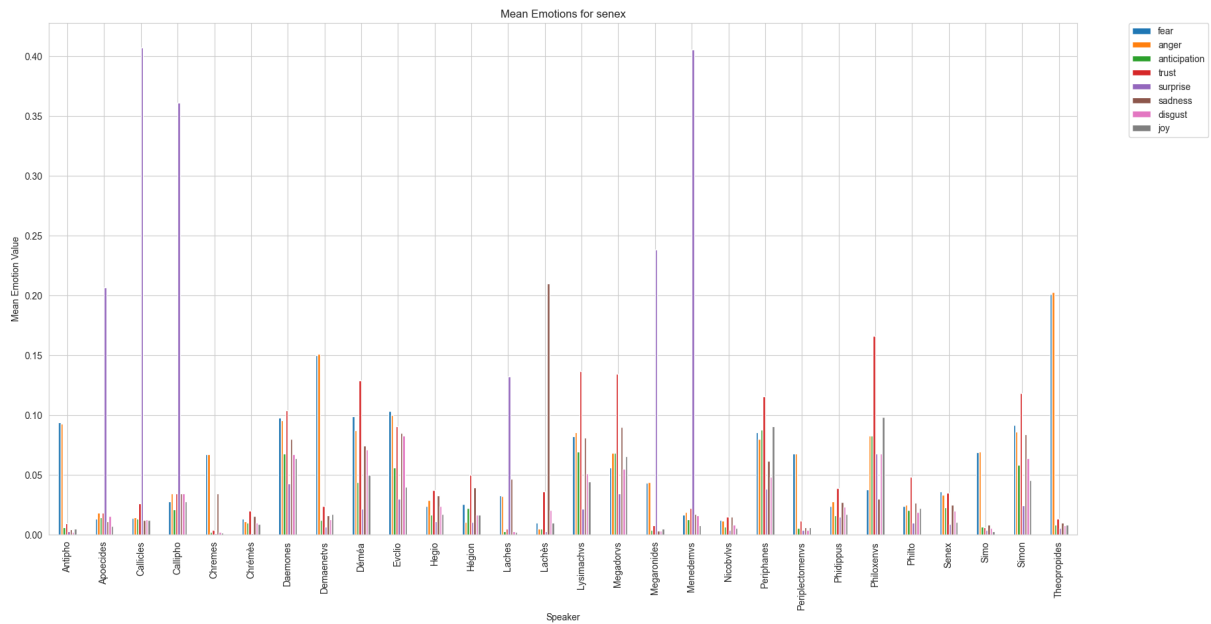


Figure 4: Emotions of *senexes* in Plautus and Terentius' drama