



**HAL**  
open science

# Linguistic Signatures for Enhanced Emotion Detection

Florian Lecourt, Madalina Croitoru, Konstantin Todorov

► **To cite this version:**

Florian Lecourt, Madalina Croitoru, Konstantin Todorov. Linguistic Signatures for Enhanced Emotion Detection. The ACM Web Conference 2026, Apr 2026, Dubai, United Arab Emirates. In press, <10.1145/3774904.3793805>. <hal-05530395v2>

**HAL Id: hal-05530395**

**<https://hal.science/hal-05530395v2>**

Submitted on 3 Mar 2026

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons CC BY-NC-ND 4.0 - Attribution - Non-commercial use - No Derivative Works - International License

# Linguistic Signatures for Enhanced Emotion Detection

Florian Lecourt  
LIRMM, Université de Montpellier  
Montpellier, France

Madalina Croitoru  
LIRMM, Université de Montpellier  
Montpellier, France

Konstantin Todorov  
LIRMM, Université de Montpellier  
Montpellier, France

## Abstract

Emotion detection is a central problem in NLP, with recent progress driven by transformer-based models trained on established datasets. However, little is known about the linguistic regularities that characterize how emotions are expressed across different corpora and labels. This study examines whether linguistic features can serve as reliable interpretable signals for emotion recognition in text. We extract emotion-specific linguistic signatures from 13 English datasets and evaluate how incorporating these features into transformer models impacts performance. Our RoBERTa-based models enriched with high level linguistic features achieve consistent performance gains of up to +2.4 macro F1 on the GoEmotions benchmark, showing that explicit lexical cues can complement neural representations and improve robustness in predicting emotion categories.

## CCS Concepts

• Computing methodologies → Natural language processing.

## Keywords

Emotion Detection, RoBERTa, GoEmotions, CARER, SEANCE, Linguistic Features

### ACM Reference Format:

Florian Lecourt, Madalina Croitoru, and Konstantin Todorov. 2026. Linguistic Signatures for Enhanced Emotion Detection. In *Proceedings of the ACM Web Conference 2026 (WWW '26)*, April 13–17, 2026, Dubai, United Arab Emirates. ACM, New York, NY, USA, 4 pages. <https://doi.org/10.1145/3774904.3793805>

## 1 Introduction

Emotion recognition in text has become increasingly relevant for applications in social media analysis, dialog systems, and affective computing [30]. Transformer-based models such as BERT [10] and RoBERTa [18] have achieved strong results on established benchmarks [1]. In parallel, linguistic tools such as LIWC [22] and SEANCE [8] capture interpretable lexical, cognitive, and affective features of language. However, explicit integration of hand-crafted or (psycho)linguistic features remains rare [40]. Moreover, probing analyses indicate that even when linguistic information is implicitly learned, many relevant features are not effectively utilized [35]. This leaves open questions about how linguistic cues relate to model predictions [17]. This study addresses this gap by connecting linguistic and neural perspectives for emotion detection. This approach is particularly relevant for Web content analysis and mining. We are guided by two main research questions (RQ):

**RQ1:** Can linguistic feature analysis across diverse benchmark datasets produce consistent and distinctive linguistic signatures for individual emotions?

**RQ2:** Do these linguistic signatures improve the performance of transformer-based emotion detection models?

To answer these questions, we extract linguistic signatures for each emotion label across thirteen publicly available English emotion datasets using SEANCE. We then integrate these features into transformer architectures leading to improved macro-F1 scores by up to +2.4 on the GoEmotions benchmark [9], and to effective generalizability, as evidenced by comparable performance to the State-of-the-art (SotA) on the CARER dataset [28]. This indicates that linguistic cues are stable transferable signals for our task.

The paper is structured as follows: Section 2 presents the dataset selection, linguistic feature extraction, and signature generation; Section 3 introduces the RoBERTa- and LLaMA-based models incorporating linguistic features; Section 4 discusses related work.

## 2 Dataset Selection and Linguistic Analysis

To build a broad basis for linguistic analysis, we conducted a systematic survey of emotion-related datasets that have shaped research in Natural Language Processing (NLP) since the Transformers paper [36]. We used the keyword **emotion** to query leading NLP venues (ACL, EMNLP, COLING, etc.) from 2017 to 2024, retrieving 129 publications. After manual screening, 39 were deemed directly relevant to emotion detection, from which we extracted the datasets used for training or evaluation. Any dataset cited in these papers was also included to ensure comprehensive coverage.

Because the vast majority of emotion datasets identified in our survey are in English, we restrict our study to English corpora. We further excluded resources that were inaccessible, non-standardized, or diverged from their documented format. As a result, 13 datasets remained, covering a wide range of textual contexts, from short social media posts to scripted dialogues, summarized in Table 1.

SEANCE is based on the General Inquirer (GI) lexicon [32], which defines more than 180 semantically interpretable linguistic categories (e.g., Positiv, Legal, Virtue). The tool outputs normalized frequency values for each category and automatically accounts for negation. Comparative evaluations suggest that SEANCE performs on par with, and in some cases slightly outperforms, LIWC in capturing affective linguistic cues [11]. SEANCE has been used in prior work for large-scale discourse analysis in various domains, demonstrating its fitness for linguistic analysis [3].

To prepare data for feature extraction, all texts or utterances in each dataset were grouped by their corresponding emotion labels, with a dedicated plain-text file created for each emotion. In multi-label datasets, texts were duplicated under every relevant label to preserve co-occurrence information. SEANCE was then applied to each emotion-specific file, generating eigenvalues that capture the distribution of linguistic features across emotions. From these outputs, we retained the 10% most frequent linguistic features per



This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.

*WWW '26, April 13–17, 2026, Dubai, United Arab Emirates*

© 2026 Copyright held by the owner/author(s).

ACM ISBN 979-8-4007-2307-0/2026/04

<https://doi.org/10.1145/3774904.3793805>

Title	Year	Source	Domain	Format	Size	# of labels
Affective Text [33]	2007	News Headlines	News	Snippets	1,250 headlines	6
BU-NEMO+ [13]	2022	News Headlines	News	Snippets	1,297 headlines	8
Cancer Emo [31]	2020	Online Cancer Survivors Network	Health	Snippets	25,000 online comments	8
CARER [28]	2018	Twitter	Social Media	Snippets	664,462 tweets	6
EmoContext [6]	2019	Conversation with Conversational Agent	Conversational Agent	Dialogues	38,424 dialogues	4
EmoProgress [37]	2024	Self-report	Psychological Self-report	Dialogues	149 dreams + 339 dialogues	10
EmoryNLP [39]	2018	Friends TV show	TV show	Dialogues	897 scenes, forming 12,606 utterances	7
Empathetic Dialogues [27]	2019	ParLAI platform	Human-Human conversation	Dialogues	24,850 conversations	32
GoEmotions [9]	2020	Reddit	Social Media	Snippets	58,009 Reddit comments	28
GoodNewsEveryone [5]	2019	News Headlines	News	Snippets	10,470 headlines	16
IDEM [26]	2024	Generated with GPT-4	Text Generation	Snippets	9,685 sentences	36
ISEAR [29]	1994	Self-reporting statements	Psychological Self-report	Snippets	7,666 texts	7
MELD [24]	2019	Friends TV show	TV show	Dialogues	1,400 dialogues forming 13,000 utterances	7

Table 1: Overview of the datasets

emotion, balancing noise reduction with coverage of salient emotional markers. These subsets define the **linguistic signatures** of emotions: compact representations of the linguistic categories most strongly associated with each label. (Cf. Figure 1 for an example.)

To enable cross-dataset comparison, emotion labels were normalized by merging semantically equivalent categories (e.g., joy and joyful), yielding a harmonized set of 30 emotions. For each, we kept only GI features present in at least 50% of instances to emphasize consistent rather than dataset-specific patterns.

We examined inter-emotion overlap using pairwise Jaccard similarity of their feature signatures. Among 435 emotion pairs, 60 showed strong overlap ( $J > 0.7$ ), revealing recurring linguistic patterns. Features such as Active\_GI (words implying active orientation; e.g., swim, take), Iav\_GI (interpretative action verbs; e.g., encourage, flatter), and Strong\_GI (strength or power; e.g., achieve, great) appeared in over 90% of emotions, whereas Virtue\_GI (moral approval or good fortune; e.g., honest, generous, unique to admiration), Hostile\_GI (hostility or aggressiveness; e.g., hate, attack, unique to anger), and Need\_GI (intent or desire; e.g., want, hedonistic, unique to desire) were emotion-specific. These linguistic signatures form a reproducible basis for cross-context emotion comparison. No strong correlation emerged between linguistic features and dataset modality (snippets vs. dialogues) or domain (news, social media, TV transcripts). The code and signatures are made available.<sup>1</sup> This analysis showed that consistent linguistic patterns can be identified across datasets (RQ1). The next section tests whether these patterns can inform transformer-based emotion models (RQ2).

### 3 Enhancing Emotion Detection Models

We adopt RoBERTa due to its slight macro-F1 advantage over other transformers in multi-label emotion tasks [7, 21]. All enhanced RoBERTa-based models were trained and evaluated on the GoEmotions corpus [9], to ensure comparability with the evaluation framework of [15], who assessed ChatGPT across various NLP tasks, including emotion detection, reporting a BERT baseline with a macro F1 of 52.75 as the SOTA. GoEmotions was additionally chosen because it is a fine-grained, multi-label dataset covering 28 emotion categories, making it a more challenging and comprehensive benchmark compared to coarse-grained, single-label datasets. To verify reproducibility, we reimplemented their BERT model with the same preprocessing pipeline, obtaining 49.67 macro F1. This serves as our reproduction baseline. We normalized text following [16]: lowercase, hashtag and emoticon handling, and slang expansion to mitigate social-media noise. Training used five random seeds (1, 2, 10, 21, 42) with early stopping (patience = 3) and AdamW ( $1 \times 10^{-5}$ ).

The first proposed model, RoBERTa-LexEnhance, is inspired by [16]. It augments RoBERTa by concatenating its [CLS] representation ( $h_{CLS}$ ) with a SEANCE-derived vector ( $s$ ) summarizing the relative frequency of each emotion’s General Inquirer (GI) categories in the input text. The final classifier input is the concatenation [ $h_{CLS}; s$ ], followed by a dropout layer (0.2) and a linear classifier.

The second proposed model, RoBERTa-EarlyFusion, integrates lexical affective features at the token level, following [38]. Each preprocessed sentence is tokenized, and every token is mapped to a binary vector encoding its associated GI categories. These token-level GI vectors are projected into an embedding space and transformed through a GELU-activated linear layer followed by a gating mechanism. The gate learns how strongly each GI feature

<sup>1</sup><https://github.com/FlorianLecourt/Linguistic-Signatures-for-Enhanced-EmotionDetection>



