（ 装 订 线 内 不 要 答 题 ）

**复旦大学外国语言文学学院**

**2024～2025学年第二学期期末考试试卷**

**🗹A卷 B卷 C卷**

**课程名称： 英语语言史 课程代码： FORE130374.01**

**开课院系： 外国语言文学学院 考试形式：开卷/闭卷/课程论文/其他**

**姓名： 杨云雅 学号： 23200120179 专业： 英语**

提示：请同学们秉持诚实守信宗旨，谨守考试纪律，摒弃考试作弊。学生如有违反学校考试纪律的行为，学校将按《复旦大学学生纪律处分条例》规定予以严肃处理。

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 题号 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 总分 |
| 得分 |  |  |  |  |  |  |  |  |  |

（以下为试卷正文**或课程论文题目**）

**A Corpus-Based Study of Emotive Adverbs Alas and Unfortunately: Syntactic Position, Register Variation, and Superlinguistic Implications**

**1. Introduction**

Recent work in formal semantics increasingly focuses on non-truth-conditional meaning, especially **within the framework of use-conditional meaning** (Cheng & Shen, 2022)**.** Under such context, *alas* and *unfortunately* are often seen as typical **emotive markers** (Rett, 2021). Their emotional content is usually not-at-issue—it cannot be easily denied and doesn’t alter the truth conditions of the sentence. This makes them similar in function to emojis or facial expressions. For example, Grosz et al. (2021) shows that in “Unfortunately, Jane lost the race,” one can reject the fact (“Jane lost”) but not the speaker’s regret. The same applies to *alas*.

While earlier studies have mainly focused on the semantics and inferential roles of *alas* and *unfortunately*, their actual usage patterns in natural discourse remain underexplored. This study aims to fill that gap by investigating **two key questions:** (1) Do *alas* and *unfortunately* differ systematically in their **syntactic position** within sentences (e.g: sentence-initial, medial, or final)? (2) Do they show different patterns of **distribution across registers** such as news, fiction, academic writing, and spoken language? And, if possible, what **insights for integrating superlinguistic elements like emojis into formal semantic analysis might it bring?**

**2. Research Method**

To answer these questions, this study uses the **9-million-word sample version of the COCA corpus** and searches for *alas* and *unfortunately* using **AntConc**. The first part of the analysis focuse on sentence position. Since *alas* is relatively rare (67 tokens) and only use as an interjection, we manually removed one example with quotation marks and annotated the rest by hand.

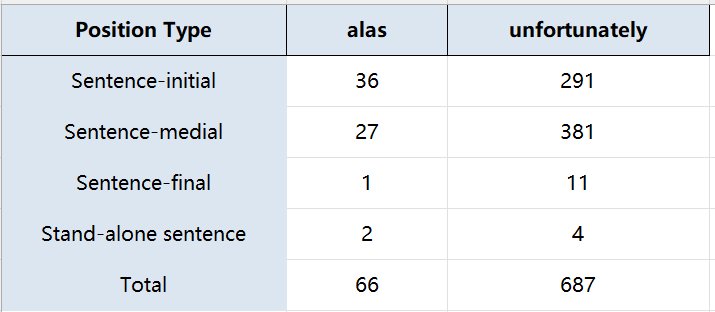
*Unfortunately* was much more complicated. Besides being an emotive marker, it often functions as a regular adverb, as in “*unfortunately*+ adj./v.” For example, “*…someone who deals with the language of terror and fear,* ***unfortunately acts*** *like these are going to make it difficult.”* To focus only on its expressive use, we filtered for tokens that were directly followed by punctuation—commas for sentence-initial and -medial use, and periods or other sentence-ending marks for sentence-final use. After filtering, there were still 387 tokens, so we use basic rules to classify them: capitalized tokens were treated as sentence-initial, those followed by punctuation as sentence-final, and the overlaps of the two mentioned above as stand-alone sentences. Others were counted as sentence-medial.

For register analysis, we simply counted how often each word appeared in files labeled by genre in the corpus (e.g., spoken, academic, blog).

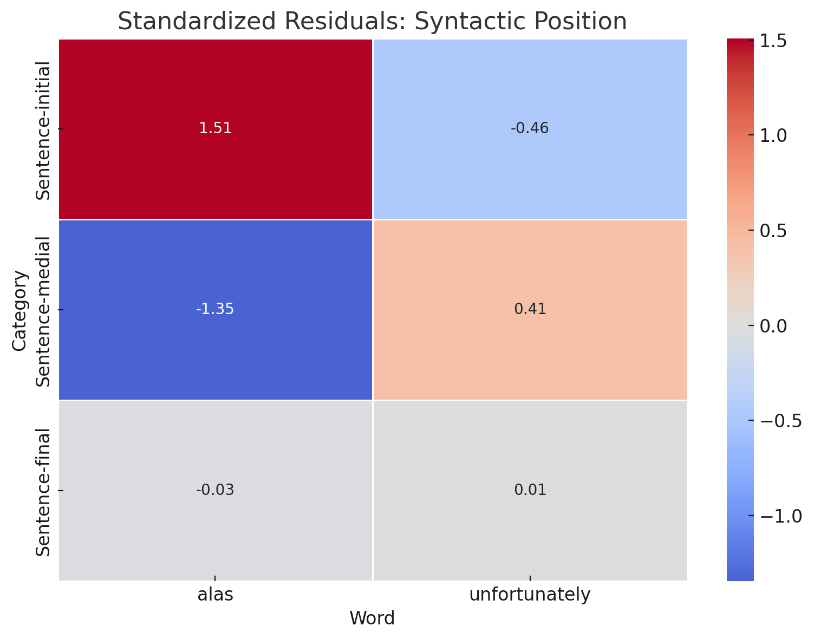
After data selection, **we apply Python on Google Colab to statistical tests and visualizations.** **Fisher’s exact test** was use for small or uneven data, while **chi-square test** was use for larger tables like registers. We use the **scipy.stats package** for testing, organized results using **Pandas**, and visualized standardized residuals with heatmaps with **Seaborn.**

**3.Findings**

**3.1 Syntactic Position**



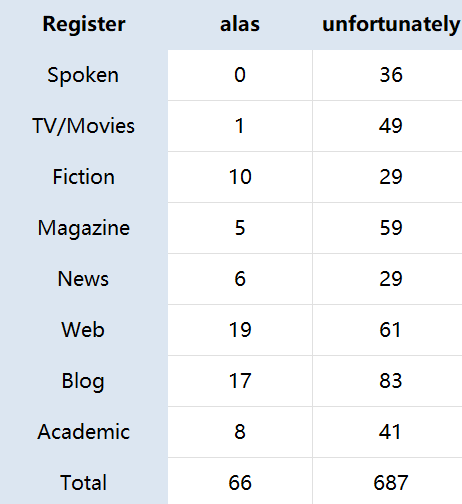
For sentence position analysis, stand-alone sentence cases were excluded. Only three main syntactic positions—sentence-initial, sentence-medial, and sentence-final—were kept to form a 2×3 contingency table. The standardized residual heatmap is shown below.



In analyzing the syntactic distribution of the two words, we noted that *alas* had a relatively small total frequency (67 tokens), with only one occurrence in sentence-final position. Given the overall low frequency and the limited variety in its positional distribution, we apply Fisher’s exact test to examine the significance of its distribution in a binary classification: sentence-initial vs. non-initial position.

The standardized residual for *alas* in the sentence-initial position was 1.51. Although this does not meet the conventional threshold for statistical significance (typically |z| > 1.96), it indicates a positive deviation—***alas* appeared in sentence-initial position more frequently than expected.** The p-value obtained from Fisher’s exact test was 0.047, indicating a statistically significant difference in the initial position distribution between *alas* and *unfortunately* (p < 0.05). This supports the observation that ***unfortunately* is more likely to occur in medial positions as a parenthetical expression, while *alas* strongly prefers the sentence-initial slot and almost never appears at the end of a sentence.** This contrast may reflect deeper grammatical and stylistic differences between the two expressions.

**3.2 Register Distribution**

*Alas* has strongly negative residuals in the Spoken and TV/Movies registers, both well below -1.96, indicating statistically significant underuse in these genres. On the other hand, **its residuals are positively skewed in Fiction and Web, with values above 1.8, suggesting an overrepresentation close to significance.** *Unfortunately*, by contrast, **shows residuals close to zero across most registers, indicating relatively balanced and widespread use, consistent with its status as a stylistically neutral adverb.**

We then use a chi-square test to assess whether the distribution of the two words was associated with register type. We also examined expected frequencies and standardized residuals to identify specific preferences or aversions in particular categories. The test yielded χ²(7) = 24.87, with a p-value of 0.0008, indicating a highly significant difference in register distribution (p < 0.001). This result statistically confirms the trends observed in the residual heatmap: *alas* is significantly overrepresented in written genres and underrepresented in speech-related ones.

Thus, we can conclude that ***alas* tends to appear in written or literary registers** such as Fiction, Academic, and Web, while avoiding more conversational or performative genres like Spoken and TV/Movies. ***Unfortunately*, on the other hand, occurs more freely across both written and spoken contexts, especially in media-related and spoken registers.**

**4. Conclusion**

It turns out that *alas* and *unfortunately*, as two adverbs use to express regret, may not behave similarly in usage when corpus analysis revealed clear the difference. *Alas* is almost absent in spoken and media-related registers, which is somewhat unexpected. But this might be due to its more “performative” nature. **My assumption is that,** **as a pure emotional signal, *alas* may be easily replaced in actual speech by non-verbal actions like sighing, much like how one cannot literally “say” an emoji during conversation.** Meanwhile, its frequent use in web-based texts suggests a closer relationship with visual or expressive elements such as emojis, particularly those use in online contexts.

By contrast, *unfortunately* is much more adaptable. **It appears flexibly across syntactic positions and is more evenly distributed across registers.** Its ability to co-exist with both propositional and expressive content may indicate that, rather than mirroring emojis directly, ***unfortunately* has more in common with co-speech gestures**—such as waving or shrugging—that accompany speech under the scale of super linguistics.

We also notice that although some existing emojis like 😔 convey similar feelings of regret, **their typical sentence position is at the end**, according to the research done by emojiall.com, **in contrast to the initial or medial preference of both *alas* and *unfortunately*.** This positional mismatch points to a possible structural difference in how emotional content is organized in linguistic versus visual-symbolic systems—a topic that calls for further investigation in future research.

**5. References**

Cheng, G., & Shen, Y. (2022). Research on the New Development of Formal Linguistics. Tsinghua University Press.

程工, & 沈园. (2022). 形式语言学新发展研究. 清华大学出版社.

Ebert, C. (2024). Semantics of Gesture. Annual Review of Linguistics, 10, 169–189.

Ebert, C. (2014). Gestures, demonstratives, and the attributive/referential distinction. Handout of a talk given at Semantics and Philosophy in Europe (SPE 7), Berlin, Germany.

Grosz, P. G., Greenberg, G., De Leon, C., & Kaiser, E. (2021). A semantics of face emoji in discourse. Manuscript. https://ling.auf.net/lingbuzz/005981

Gutzmann, D. (2013). Expressives and beyond: An introduction to varieties of use-conditional meaning. In D. Gutzmann & H.-M. Gärtner (Eds.), Beyond expressives: Explorations in use-conditional meaning (pp. 1–58). Leiden: Brill.

Rhett, J. (2021). The semantics of emotive markers and other illocutionary content. Journal of Semantics, 38(2), 305–340.