Exploring Linguistic Influence: A Comparative Analysis of Credibility Perception in German and English Educational Blog Articles

ANONYM AUTHOR

ACM Reference Format:
anonym Author. 2024. Exploring Linguistic Influence: A Comparative Analysis of Credibility Perception in German and English Educational Blog Articles. 1, 1 (April 2024), 7 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 INTRODUCTION

With the constantly growing volume of information on the Internet, users find it increasingly challenging to discern which sources they can trust [18]. Users leverage resources such as educational blog posts to gain knowledge on specific topics. In this context, credibility plays a crucial role, aiding in personalizing results and ensuring the presentation of content that is most relevant to them [7, 17]. Research also shows that users have a preference for different languages depending on the task at hand [10, 15]. To delve deeper into this phenomenon, our objective is to investigate the impact of language on the credibility of educational blog articles. Despite these insights, three critical aspects remain uncertain regarding the influence of language on credibility as induced by blog posts. Firstly, it is acknowledged that language exerts an influence on credibility in specific contexts [13], yet the extent of this effect and its generalizability on educational articles remains ambiguous. Secondly, this study aims to explore whether the language level of a text impacts its perceived credibility. The research intends to investigate the impact of language complexity on the comprehension and credibility assessment of information. In a user study, we will address blog posts in various languages (English and German) and at different levels of complexity. Subsequently, both the comprehension of the text and the assessment of the credibility of the content will be measured.

2 STUDY-DESIGN

2.1 Study type

The study employs a Latin Square Design in a within-subjects framework, encompassing four conditions, each corresponding to a distinct level of language difficulty. Each participant is presented with four different topics, with each topic being paired with a respective condition. The arrangement of conditions and topics is uniquely sequenced for each participant following the principles of the Latin Square Design.

Author's address: anonym Author.

Manuscript submitted to ACM

2.2 Research Questions

- RQ1: Do languages have an impact on the credibility of educational blog articles?
- RQ2: To what extent do differences in language proficiency between readers and authors influence the perceived credibility of content?

3 HYPOTHESES

3.0.1 *Hypothesis related to RQ1.*

- Hypothesis1 (H1): The choice of language affects the perception of credibility of educational blog articles.
  Rationale. Several indicators of credibility have been identified in the literature, including language use [9, 11]. We predict that the language presented in an educational blog post can have an impact on its perceived credibility.

3.0.2 *Hypothesis related to RQ2.*

- Hypothesis2 (H2): The perceived language proficiency of the reader influences the perception of content credibility.
  Rationale. Non-native speakers may have varying levels of language proficiency. This expectation may be reflected in the reader's comprehension and reading speed [12]. Our hypothesis is that language proficiency significantly affects the comprehensibility of blog posts, which in turn can impact their perceived credibility.
- Hypothesis3 (H3): The difficulty of the text influences the perception of content credibility.
  Rationale. Texts vary in difficulty, and their complexity can be assessed using tools such as the Flesh Reading Ease [4]. The reader's comprehension and reading speed may be affected by the text's complexity [12]. Therefore, the comprehensibility of blog posts can be directly affected, which can influence the perceived credibility of the content.
- Hypothesis4 (H4): The perceived credibility of a educational blog article is influenced by the interaction between the language proficiency of a reader and the difficulty of the text.
  Rationale. It is suggested that the perceived credibility of educational blog content is not determined solely by the reader's language proficiency (refer to H3) or the text difficulty in isolation (refer to H4), but rather by the interplay between these two factors [8].

4 METHOD

4.1 Materials

4.1.1 Data. For this study, we selected four articles from the Medium[1] platform, covering the categories of Sports, Health, Travel, and Finance, based on their popularity rankings on the platform[6]. This selection ensures that our study covers the thematic subjects typically discussed in educational blog posts. The textual content and accompanying graphics from the articles will be the exclusive focus, excluding author names, ratings, or publication institutions. This is to avoid potential bias and language distortion.
This enables us to ensure that

- Viewpoints are balanced (as each participant receives a text on the same topic)
- Knowledge bias due to diverse topics is reduced
- Control over the difficulty of the texts is gained

For this reason, the selected topics are as follows:

- 7 History's Hidden Extreme Sports That Sound Unreal [3]
- The Problem with Self-Diagnosis and Labeling Others with a Mental Health Condition[14]
- Smart Travel Hacks for Budget-Friendly Trips[16]
- 15 useful Statistics Terms to know for Finance and Investing[5]

4.1.2 Rewriting. The four articles will be rewritten into four separate texts using ChatGPT[1] and DeepL[2] covering the conditions of Easy English, Easy German, Difficult English, and Difficult German. First, we will rewrite the original article into the two English versions (easy/hard), and then we will translate the two English articles into their respective German counterparts to ensure the same information in all text variants. To verify the difficulty, the Flesch-Reading-Ease score [2, 4] will be employed. We also ensure that all texts have equal length with Reading Time of 1:30 to 3:30 minutes. The Flesch-Reading-Ease score will be summarized in such a way that a score between 30 and 100 in both German and English is considered indicative of easy readability, while scores beyond this threshold are classified as difficult. This feature allows users to view texts on different topics with varying levels of difficulty and evaluate their credibility. The text is presented in a standard blog post format, with variables such as author names and likes hidden to prevent any distortion of the results. The blog posts are presented as follows:

Fig. 1. The blog articles are presented as PDFs within a stylesheet of a blog article.

We computed the Flesch Reading Ease score by developing a Jupyter notebook capable of executing the Flesch Reading Ease formula, incorporating parameters such as the number of syllables, words, and sentences within the provided text samples. Numerical data pertaining to the text samples was derived using online tools [3]. Given the linguistic disparities between German and English, separate formulas were employed to calculate the Flesch Reading Ease scores for the respective languages.

For English texts, we utilized the formula[4]:

$$FRE = 206,835 - (1,015 * ASL) - (84,6 * ASW)$$

while for German texts, we employed the formula[2]:

$$FREger = 180 - ASL - (58,5 * ASW)$$

[1]Open AI, ChatGPT Version: 4.0, 2024 further information in section 6
[2]DeepL SE, DeepL Version 02.2024
[3]https://wordcount.com/ Version 02.2024

where ASL represents the Average Sentence Length and ASW denotes the Average Number of Syllables per Word.

The accuracy of the Flesch Reading Ease score is contingent upon the precision of syllable, word, and sentence counting. Given that text difficulty is categorized merely as "hard" or "easy," we contend that our methodology's precision suffices for our purposes.

In order to compare the texts, they must have the same word count within +/- 300 words. In addition, the Flesch Reading Ease Index should be the same within a range of +/- 30. This ensures that the texts are comparable[4].

4.2 Variables

4.2.1 Independent Variables.

- Text Readability: Easy and Difficult Readability Conditions
- Language: English and German Blog Articles

4.2.2 Dependent Variables.

- Text Credibility: The assessment of the article's credibility by the participant.

4.3 Planned sample in pre-study

All subjects will be recruited via University and will receive an expense allowance of 0.5 VP. The study should take each participant about 20 minutes to complete.

4.4 Procedure

- (1) Participants are initially provided with a brief introduction to the study. After this they will complete the informed consent process.
- (2) Participants are surveyed to evaluate their pre-existing knowledge and understanding of the study's subject matter open and closed questions. This step is crucial to ascertain their baseline familiarity with the topic.
- (3) Participants are assigned a blog post pertaining to the theme from step (2), presented in either German or English.
- (4) Following the reading of the blog post, participants undergo an evaluation to measure their level of understanding of the content. Additionally, the credibility to the information presented in the post is also assessed.
- (5) 5. Steps 2 to 4 above are repeated four times for each participant, so that they are required to evaluate four different articles.

[4]Text details are shown in the section 7

5 ANALYSIS PLAN

Given the Latin square design and four datapoints per participant, we aim to ensure statistical validity with a sample size that accommodates the complexity of our study. A power analysis conducted using G*Power software (Faul et al., 2007; Faul et al., 2009) indicates a minimum of 32 participants (multiple of 16 variants) to adequately power our results. We will conduct an exploratory analysis to investigate the hypotheses outlined in our study.

Hypotheses 1 will be examined using an paired t-tests. These test will assess whether there is a statistically significant difference in credibility assessment based on language use. We will set a significance threshold (alpha) of 0.0125 and aim for a statistical power of 0.8.

To examine Hypotheses 2, 3, and 4, we will utilize linear regression models. These models will assess the influence of language proficiency, text difficulty, and their interaction on the perceived credibility of educational blog articles. Our analysis will consider an effect size (f) of 0.16, a significance threshold (alpha) of 0.0125, and a desired statistical power of 0.8.

REFERENCES

[1] 2024. *Medium – Where good ideas find you.* https://medium.com

[2] Toni Amstad. 1978. *Wie verstaendlich sind unsere Zeitungen?* 187 S. pages.

[3] Faze. 2024. *7 History's Hidden Extreme Sports That Sound Unreal.* https://medium.com/exploring-history/7-historys-hidden-extreme-sports-that-sound-unreal-e3124c73ec30

[4] Rudolf Flesch. 1948. A new readability yardstick. *Journal of Applied Psychology* 32, 3 (1948), 221–233.

[5] C. M. F. of T. F. Fox. 2023. *15 useful Statistics Terms to know for Finance and Investing.* https://trading-data-analysis.pro/15-useful-statistics-terms-to-know-for-finance-and-investing-81350d176f12

[6] The Writers' Handbook. 2022. *The 43 biggest tags on Medium—And tips on how to use them.* https://medium.com/@writerbites/the-43-biggest-tags-on-medium-and-tips-on-how-to-use-them-a20a6bdd89fd

[7] Anikó Hannák, Piotr Sapiezynski, Arash Molavi Kakhki, Balachander Krishnamurthy, David M. J. Lazer, Alan Mislove, and Christo Wilson. 2013. Measuring personalization of web search. *Proceedings of the 22nd international conference on World Wide Web* (2013). https://api.semanticscholar.org/CorpusID:12523299

[8] Seyede Faezeh Hosseini Alast and Sasan Baleghizadeh. 2021. The interplay of glossing with text difficulty and comprehension levels. *Language Teaching Research* (2021), 13621688211020423.

[9] Juliane House. 2006. Communicative styles in English and German. *European Journal of English Studies* 10, 3 (2006), 249–267. https://doi.org/10.1080/13825570600967721 arXiv:https://doi.org/10.1080/13825570600967721

[10] Mitesh M. Khapra, Salil Rajeev Joshi, Ananthakrishnan Ramanathan, and Karthik Venkat Ramanan. 2013. Offering language based services on social media by identifying user's preferred language(s) from romanized text. *Proceedings of the 22nd International Conference on World Wide Web* (2013). https://api.semanticscholar.org/CorpusID:10182124

[11] Shiri Lev-Ari and Boaz Keysar. 2010. Why don't we believe non-native speakers? The influence of accent on credibility. *Journal of Experimental Social Psychology* 46 (2010), 1093–1096. https://api.semanticscholar.org/CorpusID:144696826

[12] Jessica T Maluch and Karoline A Sachse. 2020. Reading in Developing L2 Learners: The Interrelated Factors of Speed, Comprehension and Efficiency across Proficiency Levels. *Tesl-Ej* 24, 1 (2020), n1.

[13] Florian Maximilian Meier. 2021. The Influence of Multilingualism and Mutual Intelligibility on Wikipedia Reading Behaviour: A Research Proposal. In *Information between Data and Knowledge. Information Science and its Neighbors from Data Science to Digital Humanities*, T. Schmidt and C. Wolff (Eds.), Vol. 74. Wallstein Verlag, 64–72. https://doi.org/10.5283/epub.44937 The 16 th International Symposium of Information Science, ISI 2021 ; Conference date: 08-05-2021 Through 10-05-2021.

[14] G. P. Ph.D. 2024. *The Problem with Self-Diagnosis and Labeling Others with a Mental Health Condition.* https://medium.com/wise-well/the-problem-with-self-diagnosis-and-labeling-others-with-a-mental-health-condition-785044b4d680

[15] Denise D. Quigley, Marc N. Elliott, Katrin Hambarsoomian, Shondelle M. Wilson-Frederick, William G. Lehrman, Denis Agniel, Judy H. Ng, Elizabeth H. Goldstein, Laura A. Giordano, and Steven C. Martino. 2019. Inpatient care experiences differ by preferred language within racial/ethnic groups. *Health Services Research* 54 (2019), 263 – 274. https://api.semanticscholar.org/CorpusID:58587023

[16] L. R. 2024. *Smart Travel Hacks for Budget-Friendly Trips.* https://medium.com/digital-global-traveler/smart-travel-hacks-for-budget-friendly-trips-25ff6806b2da

[17] Filip Radlinski and Susan T. Dumais. 2006. Improving personalized web search using result diversification. *Proceedings of the 29th annual international ACM SIGIR conference on Research and development in information retrieval* (2006). https://api.semanticscholar.org/CorpusID:1001779

[18] Julia Schwarz and Meredith Morris. 2011. Augmenting Web Pages and Search Results to Support Credibility Assessment. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Vancouver, BC, Canada) *(CHI '11)*. Association for Computing Machinery, New York, NY, USA, 1245–1254. https://doi.org/10.1145/1978942.1979127

6 EDITING PROMPTS FOR GPT-4

6.0.1 English Easy Prompt. "Can you take this block article and make it easier to read/understand? Substitute difficult words for easier ones:"

6.0.2 English Hard Prompt. "Can you take the following blog article and rewrite it to be for a very well read academic audience?"

6.0.3 German. As stated in the method part already, we used deepl for the german text variants.

7 TEXTDETAILS

7.1 Text Variants:

Table 1. Topic Health

Difficulty	Flesch Reading Ease	Word Count	Reading Time in Minutes
english-hard	2	512	2:02
english-easy	47	762	3:02
german-hard	4	836	3:20
german-easy	45	511	2:02

Table 2. Topic Finance

Difficulty	Flesch Reading Ease	Word Count	Reading Time in Minutes
english-hard	11	671	2:41
english-easy	70	624	2:29
german-hard	4	836	3:20
german-easy	24	723	2:53

Table 3. Topic Sports

Difficulty	Flesch Reading Ease	Word Count	Reading Time in Minutes
english-hard	16	751	2:57
english-easy	76	504	3:00
german-hard	17	753	3:00
german-easy	59	509	2:06

Table 4. Topic Travel

Difficulty	Flesch Reading Ease	Word Count	Reading Time in Minutes
english-hard	4	530	2:07
english-easy	61	451	1:48
german-hard	26	575	2:18
german-easy	53	454	1:48